(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 December 2002 (05.12.2002)

PCT

(10) International Publication Number
WO 02/096467 A2

(51) International Patent Classification[7]: A61K 47/48

(21) International Application Number: PCT/GB02/02384

(22) International Filing Date: 21 May 2002 (21.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0112687.9 24 May 2001 (24.05.2001) GB

(71) Applicant *(for all designated States except US)*: **MICROBIOLOGICAL RESEARCH AUTHORITY** [GB/GB]; CAMR, Porton Down, Salisbury, Wiltshire SP4 0JG (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SUTTON, John, Mark** [GB/GB]; Microbiological Research Authority, CAMR, Porton Down, Salisbury, Wiltshire SP4 0JG (GB). **SHONE, Clifford, Charles** [GB/GB]; Microbiological Research Authority, CAMR, Porton Down, Salisbury, Wiltshire SP4 0JG (GB).

(74) Agents: **SCHLICH, George, William** et al.; Mathys & Squire, 100 Gray's Inn Road, London WC1X 8AL (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/096467 A2

(54) **Title:** PHARMACEUTICAL USE FOR SECRETED BACTERIAL EFFECTOR PROTEINS

(57) **Abstract:** A polypeptide conjugate contains a bacterial injectable effector protein, secreted by a modified pilus or "needle-like" structure comprising a type III or type IV secretion apparatus, and a carrier that targets the conjugate to a target cell. The effector protein is used for a variety of purposes including treatment of neurodegenerative disease, intracellular infection and diseases associated with defects of secretion.

PHARMACEUTICAL USE OF SECRETED BACTERIAL EFFECTOR PROTEINS

The present invention relates to pharmaceutical use of secreted, injected bacterial effector proteins. In particular, the present invention relates to manufacture and use of such proteins and combination and conjugation of the proteins with carriers.

A number of deficiencies exist in the availability and suitability of neuronal therapies. At the present time, a large number of neuronal disorders have inadequate provisions for therapeutic intervention. For example there is currently no effective treatment for neuronal damage caused by ischemia or trauma. Other neurodegenerative disorders such as Motor neurone disease, Alzheimer's disease, Parkinson's disease and prion disorders such as CJD are all poorly addressed by current therapies. This reflects in part the complexity of the nervous system and the difficulties in targeting suitable therapies to the specific cells affected. Neuronal repair after damage is another disorder for which there is no effective treatment.

A number of neurological disorders are known that arise from neuronal trauma that stimulates nerve damage due to internal processes such as apoptosis. It is known to treat such disorders using a superoxide dismutase in combination with a components that targets the enzyme to neurons. However, further active compounds for treatment of neuronal disease are desired.

It is known to use type III effectors in pharmaceutical compositions.

US 5972899 describes a composition comprising *Shigella* IpaB, an IpaB fusion protein or a functional derivative or antagonist, or IpaB DNA for delivery to a eukaryotic cell to induce or to inhibit apoptosis. Site-specific delivery may be achieved within a targeted immunoliposome. Cell-type specificity is achieved by the incorporation of a cell-type selective monoclonal antibody into the lipid bilayer. Disadvantages associated with this delivery method include the very large size, low stability and poor tissue penetration of immunoliposomes, and difficulties associated with consistent immunoliposome manufacture for therapeutic use. There is also the likelihood of a high background effect due

to fusion of immunoliposomes with non-target cell types, caused by the inherent properties of the liposome membrane.

WO 01/19393 describes Type III effector proteins linked to a protein transduction domain of the HIV TAT protein. DNA constructs encoding the effector-transducer fusion protein are targeted to host cells comprising a Type III secretion system using a tissue-specific viral or plasmid vector. Upon expression in the transformed host cells, the effector-transducer conjugate is secreted and undergoes secondary redistribution and uptake by neighbouring cells.

The HIV TAT transduction domain is not specific to any cell type, hence, targeting of effector is carried out solely at the DNA level. Disadvantages of targeting effector DNA (rather than targeting effector protein) include the time lag for processing of effector DNA to effector protein. Where viral vectors are used, there are the risks of immunogenic effects and of the vector integrating into the genome.

WO 00/37493 describes *Bordetella pertussis* effector virulence genes associated with a Type III secretion system. The pathogenicity genes or encoded polypeptides are used in vaccine compositions and may be conjugated to another molecule or provided with a carrier for delivery. Pathogenicity polypeptide may be delivered via a vector directing expression of *Bordetella* pathogenicity polynucleotide in vivo.

WO 98/56817 describes pharmaceutical compositions comprising a non-pathogenic organism expressing the YopJ protein, and YopJ protein combined with a carrier, for delivery of YopJ to gastrointestinal cells from the gut. The delivery mechanism disclosed in this document is via the normal bacterial Type III secretion system - that is, one step from bacterium to target cell.

WO 99/52563 describes targeting of proteins produced by recombinant *Yersinia* to the cytosol of eukaryotic cells for diagnostic/ therapeutic purposes. Fusion proteins with the YopE targeting signal are expressed in *Yersinia* cells and delivered directly to eukaryotic cells via the Type III secretion system in the presence of the SycE chaperone.

US 5965381 describes the *in vitro* use of recombinant *Yersinia* to deliver proteins to eukaryotic cells for immune diagnostic and therapeutic purposes. The proteins are fused to a delivery sequence, recognised by the *Yersinia* Type III secretion system.

It is not advantageous to make use of bacteria for delivering therapeutic proteins due to the risk of iliciting an unwanted immune response.

The present invention has as an object the provision of new pharmaceutical compositions for a variety of uses. A further object is to provide new pharmaceutical compositions for treatment of neuronal cells.

Accordingly, the present invention provides new therapies based upon a new class of bacterial-derived proteins, though the scope of the invention is intended to embrace also fragments and derivatives and modifications thereof that retain the properties of the native proteins.

A first aspect of the invention thus lies in a pharmaceutical composition, comprising a bacterial injected effector secreted by the type III or IV secretion pathway.

The pharmaceutical composition can be used for treatment of a subpopulation of cells in a patient, especially for a treatment selected from promoting survival of cells, preventing damage to cells, reversing damage to cells, promoting growth of cells, inhibiting apoptosis, inhibiting release of an inflammatory mediator from cells and promoting division of cells, or for a treatment selected from inhibiting survival of cells, inhibiting growth of cells, inhibiting division of cells, promoting apoptosis, killing cells, promoting release of an inflammatory mediator from cells and regulating nitric oxide release from cells.

A carrier can be provided to target the effector protein to a target cell, optionally targeting the effector to a cell selected from an epithelial cell, a neuronal cell, a secretory cell, an immunological cell, an endocrine cell, an inflammatory cell, an exocrine cell, a bone cell and a cell of the cardiovascular system.

Another means of delivery of the effector is via a conjugate of the effector protein and the carrier, the two suitably linked by a linker. One particularly

preferred linker is cleavable, in that it can be cleaved after entry into the target cell so as to release the effector from the carrier. This linker can be a di-sulphide bridge or a peptide sequence including a site for a protease found in the target cell. In another embodiment of the invention, the linker is composed of two cooperating proteins, a first cooperating protein associated with the effector and the second associated with the cell targetting component. These respective parts can be administered separately and combine *in vivo* to link the effector to the cell targetting component. An example of such a two-part linker is botulinum toxin $C2_1$ in cooperation with $C2_2$.

In one embodiment of the invention, described in more detail below, a composition comprises a neuronal cell targeting component, linked by a cleavable linker to the effector protein. Preferably, the neuronal cell targeting component comprises a first domain targeting the effector to a neuronal cell and a second domain that translocates the effector into the cytosol of the neuronal cell.

Preparation of the compositions of the invention can be by combining a type III effector protein with a pharmaceutically acceptable carrier. In such compositions, the effector protein may be on its own or may be chemically linked with a (targetting) carrier. Another preparation method is to express a DNA that encodes a polypeptide having a first region that corresponds to the effector protein and a second region that codes for the carrier. A third region, between the first and second regions, which is cleaved by a proteolytic enzyme present in the target cell is optionally included.

A specific composition of the invention, for delivery of a bacterial type III effector protein to neuronal cells, comprises:-

the effector protein; linked by a cleavable linker to

a neuronal cell targeting component, comprising a first domain that binds to a neuronal cell and a second domain that translocates the effector protein of the composition into the neuronal cell. It is preferred that the first domain is selected from (a) neuronal cell binding domains of clostridial toxins; and (b) fragments, variants and derivatives of the domains in (a) that substantially retain the neuronal cell binding activity of the domains of (a).

It is further preferred that the second domain is selected from (a) domains of clostridial neurotoxins that translocate polypeptide sequences into cells, and

(b) fragments, variants and derivatives of the domains of (a) that substantially retain the translocating activity of the domains of (a).

In use of a composition of the invention for treatment of a neuronal condition, the linker is cleaved in the neuronal cell so as to release the effector protein from the targeting component, thus enabling the effector to have effect in the cell without being hindered by attachment to the targeting component.

Hence, also, the invention provides a method of delivering a bacterial type III effector protein to a neuronal cell comprising administering a composition of the invention.

The first domain may suitably be selected from (a) neuronal cell binding domains of clostridial toxins; and (b) fragments, variants and derivatives of the domains in (a) that substantially retain the neuronal cell binding activity of the domains of (a). The second domain is suitably selected from (a) domains of clostridial neurotoxins that translocate polypeptide sequences into cells, and (b) fragments, variants and derivatives of the domains of (a) that substantially retain the translocating activity of the domains of (a). The second domain is further suitably selected from:-

(a) a translocation domain that is not a H_N domain of a clostridial toxin and is not a fragment or derivative of a H_N domain of a clostridial toxin;

(b) a non-aggregating translocation domain as measured by size in physiological buffers;

(c) a H_N domain of a diphtheria toxin,

(d) a fragment or derivative of (c) that substantially retains the translocating activity of the H_N domain of a diphtheria toxin,

(e) a fusogenic peptide,

(f) a membrane disrupting peptide, and

(g) translocating fragments and derivatives of (e) and (f).

In an embodiment of the invention a construct comprises effector protein linked by a disulphide bridge to a neuronal cell targeting component comprising a first domain that binds to a neuronal cell and a second domain that translocates the effector protein into the neuronal cell. This construct is made recombinantly as a single polypeptide having a cysteine residue on the effector protein which forms a disulphide bridge with a cysteine residue on the second domain. The

effector protein is covalently linked, initially, to the second domain. Following expression of this single polypeptide, effector protein is cleaved from the second domain leaving the effector protein linked only by the disulphide bridge to the rest of the construct.

Particular aspects of the invention reside in further choices for the binding and translocation domains, and one such aspect provides a non-toxic polypeptide, for delivery of the effector protein to a neuronal cell, comprising:-

a binding domain that binds to the neuronal cell, and

a translocation domain that translocates the effector protein into the neuronal cell,

wherein the translocation domain is not a H_N domain of a clostridial neurotoxin and is not a fragment or derivative of a H_N domain of a clostridial toxin.

The binding domain is suitably comprised of or derived from clostridial heavy chain fragments or modified clostridial heavy chain fragments. As used herein, the term "modified clostridial heavy chain fragment" means a polypeptide fragment that retains similar biological functions to the corresponding heavy chain of a botulinum or tetanus neurotoxin but differs in its amino acid sequence and other properties compared to the corresponding heavy chain. The invention more specifically provides such constructs that are based on fragments derived from botulinum and tetanus neurotoxins.

In a further aspect, the invention also provides a polypeptide, for delivery of a effector protein to a neuronal cell, comprising:-

a binding domain that binds to the neuronal cell, and

a translocation domain that translocates the effector protein into the neuronal cell,

wherein the resulting construct is non-aggregating.

Whether the construct is an aggregating one is usually apparent from a lack of solubility of the construct, and this may be seen upon simple visual inspection of the construct in aqueous media: non-aggregating domains result in constructs of the invention that are partially or preferably totally soluble whereas aggregating domains result in non-soluble aggregates of polypeptides having apparent sizes of many tens or even hundreds the size of a single polypeptide. Generally, the construct should be non-aggregating as measured

by its size on gel electrophoresis, and domain sizes or apparent domain sizes thus measured should preferably be less than 1.0×10^6 daltons, more preferably less than 3.0×10^5 daltons, with the measuring being suitably carried out on native PAGE using physiological conditions.

A still further aspect of the invention provides a polypeptide, for delivery of a effector protein to a neuronal cell, comprising:-

a binding domain that binds to the neuronal cell, and

a translocation domain that translocates the effector protein into the neuronal cell,

wherein the translocation domain is selected from (1) a H_N domain of a diphtheria toxin, (2) a fragment or derivative of (1) that substantially retains the translocating activity of the H_N domain of a diphtheria toxin, (3) a fusogenic peptide, (4) a membrane disrupting peptide, (5) a H_N from botulinum toxin C_2 and (6) translocating fragments and derivatives of (3), (4) and (5).

It is to be noted that botulinum toxin C_2 is not a neurotoxin as it has no neuronal specificity, instead it is an enterotoxin and suitable for use in the invention to provide a non-aggregating translocation domain.

A yet further aspect of the invention provides a polypeptide, for delivery of a effector protein to a neuronal cell, comprising:-

a binding domain that binds to the neuronal cell, and

a translocation domain that translocates the effector protein into the neuronal cell,

wherein the polypeptide has reduced affinity to neutralising antibodies to tetanus toxin compared with the affinity to such antibodies of native tetanus toxin heavy chain.

The above aspects may singly or in any combination be exhibited by polypeptides of the invention and thus a typical preferred polypeptide of the invention (i) lacks the neurotoxic activities of botulinum and tetanus toxins, (ii) displays high affinity to neuronal cells corresponding to the affinity of a clostridial neurotoxin for those cells, (iii) contains a domain which can effect translocation across cell membranes, and (iv) occurs in a less aggregated state than the corresponding heavy chain from botulinum or tetanus toxin in physiological buffers.

A significant advantage of the polypeptides of particular aspects of the invention is their non-aggregated state, thus rendering them more usable as soluble polypeptides. The polypeptides according to the invention generally include sequences from the H_c domains of the botulinum and tetanus neurotoxins and these are combined with functional domains from other proteins, such that the essential functions of the native heavy chains are retained. Thus, for example, the H_c domain of botulinum type F neurotoxin is fused to the translocation domain derived from diphtheria toxin to give modified clostridial heavy chain fragment. Surprisingly, such polypeptides are more useful as constructs for delivering substances to neuronal cells than are the native clostridial heavy chains.

The current invention provides constructs containing type III secreted effector proteins and optionally other functional domains that effect the specific delivery of the type III effector moiety to neuronal cells These constructs have a variety of clinical uses for the treatment of neuronal diseases.

The type III secretion mechanism of Gram negative bacterial pathogens is a complex system used to deliver proteins to eukaryotic cells. The secretion mechanism utilises at least 10 -15 essential proteins to form an injection needle that extends from the surface of the bacteria and penetrates into the host cell. The effector proteins are then trafficked across the bacterial and host membranes through the lumen of the needle and injected directly into the cell cytosol. This process involves a still undefined secretion signal and involves specific chaperone proteins that deliver the effector to the secretion machinery. The system delivers a wide range of protein effectors capable of modulating host cell function in such a way as to allow the persistence or spread of the pathogen in the host. These effectors modulate a number of signalling pathways and one pathogen may export several effectors that regulate different pathways either concurrently or during different phases of its life cycle. Type III secretion systems have been described in a wide range of pathogenic bacteria including but not restricted to :

Mammalian pathogens; *Yersinia* species (including *pestis, pseudotuberculosis, enterocolitica*) , *Salmonella* species (including *typhimurium, enterica, dublin, typhi*) *Escherichia coli, Shigella* species (e.g *flexneri*), *Pseudomonas aeruginosa, Chlamydia* species (e.g.*pneumoniae, trachomatis*), and *Bordetella* species, and *Burkholderia* species

Plant pathogens; *Pseudomonas syringae, Erwinia* species, *Xanthomonas* species, *Ralstonia solanacearum,* and *Rhizobium* species

Insect pathogens; *Sodalis glossinidius*, *Edwardsiella ictaluri*, and *Plesiomonas* species

Effector proteins from any of these species, whether mammalian pathogens or not, have therapeutic potential for treating human or animal disease.

Table 1 lists a number of type III effectors that have been identified to date.

The type IV secretion system shows a remarkable degree of similarity to the type III system in that it forms a needle-like structure through which effector proteins are injected into the host cell cytoplasm. However, the proteins involved in the structure of the needle are different for the two systems and the effectors are also divergent. The effectors function to modulate cellular signalling to establish and maintain the intracellular niche and/or promote invasion and proliferation. The system is described as essential in a number of important bacterial pathogens including *Legionella pneumophila, Bordetella pertussis, Actinobacillus actinomycetemcomitans, Bartonella henselae, Escherichia coli, Helicobacter pylori, Coxiella burnetii, Brucella abortus, Neisseria* species and *Rickettsia* species (e.g. *prowazekii*). Similar type IV secretion systems exist in plant or invertebrate pathogens and are also a source of therapeutic agents. A number of described type IV effectors are also listed in table 1 with proposed functions.

The function of a variety of type III effectors has been described in recent years. Interestingly a number of effectors from different organisms have evolved to target particular signalling pathways suggesting some similarities in the mechanism of pathogenicity. The precise specificity of particular effectors may vary according to pathogen and cell type and this range of activities make them attractive candidates for therapeutic applications. Examples of some of the families of effectors useful in the present invention are described below:

GTPase activating proteins. YopE from *Yersinia pseudotuberculosis*, SptP from *Salmonella typhimurium* and ExoS and ExoT from *Pseudomonas aeruginosa* are all GTPase activating proteins (GAPs) for Rho family GTPases and are

characterised by a conserved "arginine finger" domain (Black and Bliska, (2000) *Molecular Microbiology* **37**:515-527; Fu and Galan (1999) *Nature* **401**:293-297; Goehring *et al* (1999) *Journal of Biological Chemistry* **274**:36369-36372). By increasing the hydrolysis of bound GTP they promote the formation of the inactive GDP-bound of the GTPase. This acts to down-regulate the function of a range of GTPases in cells. YopE is a 23kDa effector which is translocated into the cytosol of cells during infection by *Y.pseudotuberculosis* and other strains. Studies *in vitro* have shown that it acts as a GAP for RhoA, Cdc42 and Rac1, but not for Ras (Black and Bliska, (2000) *Molecular Microbiology* **37**:515-527). A point mutation within the arginine finger motif causes a loss of GAP activity and this correlates directly with its biological activity in cells. In *in vivo* studies carried out using a cell model that mimics the normal site of *Yersinia* infection YopE appears to have a greater specificity for Cdc42 (Andor *et al* (2001) *Cellular Microbiology* **3**:301-310). The GAP activity of SptP shows greater specificity for Cdc42 and Rac1 compared to RhoA. The GAP activity of particular proteins is likely to vary for different cell types and delivery routes. SptP, ExoS and ExoT are bifunctional enzymes with additional enzymatic domains (SptP, tyrosine phosphatase; ExoS, ExoT, ADP-ribosyltransferase). In the case of ExoS this activity blocks the activation of Ras GTPase allowing a co-ordinated modulation of different signalling pathways (Henriksson *et al* (2000) *Biochemical Journal* **347**:217-222).

Guanine nucleotide exchange factor. SopE and SopE2 from *Salmonella typhimurium* and related proteins act as guanine nucleotide exchange factors (GEFs) for a range of GTPases (Hardt *et al* (1998) *Cell* **93**:815). GEFs function by enhancing the rate of replacement of bound GDP by GTP causing the activation of the GTPase. This effectively upregulates the activity of specific GTPases in the cell. Native SopE is a 240 amino acid protein which is injected into the host cell cytosol by *S.typhimurium*. The N-terminal 77 amino acids of the protein act as a secretion signal and are dispensable for the biological activity of the protein (Hardt *et al* (1998) *Cell* **93**:815). In *in vitro* studies SopE acts as a GEF for CDc42, Rac1, Rac2, RhoA, and RhoG. Cellular GEFs show a high degree of specificity for particular GTPases and it is likely that SopE will show greater specificity *in vivo*. This specificity is likely to vary according to cell type and delivery route. The type IV effector, RalF, from Legionella pneumophila is a further exchange factor affecting the function of small GTPases. In this case the target is the ADP ribosylation factor (ARF) family and

this is the first example of a bacterial effector that targets this family (Nagai *et al* (2002) *Science* **295**;679-682).

Covalent modification of GTPase. The type III effector YopT from *Y.pestis* and certain other *Yersinia* strains has similar effects *in vivo* to YopE (Iriarte and Cornelis (1998) *Molecular Microbiology* **29**:915-929). In HeLa cells YopT causes a shift in the electrophoretic mobility of RhoA but not Cdc-42 or Rac (Zumbihl *et al* (1999) *Journal of Biological Chemistry* **274**:29289-29293). It is still not known whether this represents a direct modification of RhoA by YopT or whether other cellular factors are involved. The specificity of YopT for RhoA offers significant therapeutic possibilities.

Regulation of cell signalling via protein kinase and phosphatase. YopO/YpkA from *Yersinia* spp are protein kinase related to eukaryotic serine/threonine kinases (Galyov *et al* (1993) *Nature* **361**:730-732). YopO/YpkA causes a similar cell rounding to that observed for other effectors such as YopE suggesting a role in modulating GTPase function. The small GTPases RhoA and Racl have been shown to bind to YopO and YpkA suggesting that these are the intracellular targets for the kinase (Barz C *et al* (2000) *FEBS Letters* **482**:139-143). The type IV effector CagA from *Helicobacter pylori* also affects the cytoskeleton of infected cells and its activity is dependent on its phosphorylation by intracellular kinases. CagA functions via the SHP-2 tyrosine phosphatase to modulate downstream signalling.

Inositol phosphatases. SigD from *Salmonella typhimurium*, SopB from *S.dublin* and IpgD from *Shigella flexneri* are all putative inositol phosphatases. In intestinal cells SopB causes an accumulation of inositol 1,4,5,6, tetrakisphosphate. Mutations in active site of SopB abolishes both its phopshatase activity and the accumulation of inositol tetrakisphosphate (Norris *et al* (1998) *Proceedings of the National Academy of Science U.S.A* **95**:14057-14059). SopB appears to hydrolyse a wide range of inositol and phosphatidylinositol phosphates *in vitro* although its precise intracellular target remains to be defined (Eckmann *et al* (1997) *Proceedings of the National Academy of Science U.S.A* **94**:14456-14460). SigD appears to have a different specificity *in vivo* as it does not lead to an increase in the levels of inositol 1,4,5,6, tetrakisphosphate (Eckmann *et al* (1997)). Although again the precise intracellular target has not been defined, SigD has been shown to lead to the

activation of Akt /Protein kinase B in epithelial cells (Steele-Mortimer (2000) *Journal of Biological Chemistry* **275**:37718-37724). The activity has been shown to be dependent on the presence of a synaptojanin-homologous region close to the C-terminus of the protein (Marcus *et al* (2001) *FEBS letters* **494**:201-207). The homologous protein IpgD also stimulates the activation of Akt in these cells (Marcus *et al* (2001)). The potential to activate Akt offers a number of therapeutic opportunities as it is a key regulator of cellular survival (reviewed in Vanhaesebroeck and Alessi (2000) *Biochemical Journal* **346**:561-576).

Inhibition of mitogen-activated protein kinase kinase. YopJ from *Yersinia pestis* is another translocated effector with a wide range of homologs including AvrA from *Salmonella* spp. and a variety of effectors from plant pathogens. YopJ has been shown to inactivate a broad range of mitogen-activated protein kinase kinases (MKKs) (Orth *et al* (1999) *Science* **285**:1920-1923) causing apoptosis in macrophages. YopJ is suggested to act as a ubiquitin-like protein protease causing increased turnover of signalling molecules via removal of a Sumo-1 tag from the MKK (Orth *et al* (2000) *Science* **290**:1594-1597). Interestingly in cell models of cytokine production and macrophage killing AvrA shows no activity despite its homology to YopJ suggesting that the specificity of the proteins may be different (Schesser K *et al* (2000) *Microbial Pathogenesis* **28**:59-70). In neuronal cells these different specificities may offer potential therapeutic applications for modulating MKKs involved in apoptosis or inflammatory responses.

Modulators of cellular trafficking. SpiC from *Salmonella enterica* inhibits the fusion of endosomal vesicles to prevent the exposure of *Salmonella* to lyosomal degradation (Uchiya *et al* (1999) *EMBO Journal* **18**:3924-3933). The ability to modulate intracellular trafficking pathways offers a number of therapeutic opportunities for modulating cycling of receptors or release of material from membrane bound vesicles.

A number of additional effector proteins are implicated in regulating and maintaining the intracellular compartments occupied by bacterial pathogens. *Salmonella* , in common with many other pathogens, establishes a specialised intracellular compartments. *Salmonella* has a dedicated type III secretion system that secretes proteins into the host cell cytosol from within this

compartment and the effectors secreted by this system (including SpiC, SopE/E2, SseE,F,G,J, PipA,B, SifA,B)maintain the integrity of this compartment. A recent paper described the synergistic effects of SseJ and SifA in regulating processes from the vacuolar membrane (Ruiz-Albert *et al* (2002) *Molecular microbiology* **44**;p645-661). These proteins and their counterparts from other intracellular pathogens have significant potential for treating disorders affecting intracellular trafficking pathways. RalF and a number of the other effectors described previously may also have significant therapeutic potential for such disorders.

The botulinum neurotoxins are a family of seven structurally similar, yet antigenically different, protein toxins whose primary site of action is the neuromuscular junction where they block the release of the transmitter acetylcholine. The action of these toxins on the peripheral nervous system of man and animals results in the syndrome botulism, which is characterised by widespread flaccid muscular paralysis (Shone (1986) in 'Natural Toxicants in Foods', Editor D. Watson, Ellis Harwood, UK). Each of the botulinum neurotoxins consist of two disulphide-linked subunits; a 100 kDa heavy subunit which plays a role in the initial binding and internalisation of the neurotoxin into the nerve ending (Dolly et. al. (1984) Nature, 307, 457-460) and a 50 kDa light subunit which acts intracellularly to block the exocytosis process (McInnes and Dolly (1990) Febs Lett., 261, 323-326; de Paiva and Dolly (1990) Febs Lett., 277, 171-174). Thus it is the heavy chains of the botulinum neurotoxins that impart their remarkable neuronal specificity.

Tetanus toxin is structurally very similar to botulinum neurotoxins but its primary site of action is the central nervous system where it blocks the release of inhibitory neurotransmitters from central synapses (Renshaw cells). As described for the botulinum toxins above, it is domains within the heavy chain of tetanus toxin that bind to receptors on neuronal cells.

The binding and internalisation (translocation) functions of the clostridial neurotoxin (tetanus and botulinum) heavy chains can be assigned to at least two domains within their structures. The initial binding step is energy-independent and appears to be mediated by one or more domains within the H_c fragment of the neurotoxin (C-terminal fragment of approximately 50kDa) (Shone *et al.* (1985), Eur. J. Biochem., 151, 75-82) while the translocation step

is energy-dependent and appears to be mediated by one or more domains within the H_N fragment of the neurotoxin (N-terminal fragment of approximately 50kDa).

Isolated heavy chains are non-toxic compared to the native neurotoxins and yet retain the high affinity binding for neuronal cells. Tetanus and the botulinum neurotoxins from most of the seven serotypes, together with their derived heavy chains, have been shown to bind a wide variety of neuronal cell types with high affinities in the nM range (e.g botulinum type B neurotoxin; Evans *et al.* (1986) Eur. J. Biochem. 154, 409-416). Another key characteristic of the binding of the tetanus and botulinum heavy chains to neuronal cells is that the receptor ligand recognised by the various toxin serotypes differ. Thus for example, botulinum type A heavy chain binds to a different receptor to botulinum type F heavy chain and these two ligands are non-competitive with respect to their binding to neuronal cells (Wadsworth *et al.* (1990), Biochem J. 268, 123-128). Of the clostridial neurotoxin serotypes so far characterised (tetanus, botulinum A, B, C_1, D, E and F), all appear to recognise distinct receptor populations on neuronal cells. Collectively, the clostridial neurotoxin heavy chains provide high affinity binding ligands that recognise a whole family of receptors that are specific to neuronal cells.

The present invention also provides constructs for the delivery of type III effector proteins specifically to neuronal cells. The mechanism by which the type III effector protein is delivered to the cell by these constructs is completely different to that used by the host bacteria. Instead of being injected directly into the cellular cytosol, specific constructs of the invention deliver the type III effector protein to cells via a number of sequentially acting biologically active domains and by a process resembling receptor-mediated endocytosis. Surprisingly, when delivered by this completely different mechanism, the type III effector proteins are biologically active within the cellular cytosol.

Particular constructs of the invention comprise three functional domains defined by their biological activities. These are:

the type III effector moiety (for examples see Table1);

a targeting domain that binds the construct to receptors and that provides a high degree of specificity to neuronal cells; and

a translocation domain that after internalisation of the construct, effects

the translocation of the type III effector moiety through the endosomal membrane into the cell cytosol.

The type III effector-containing construct may also contain 'linker proteins' by which these domains are interconnected. In one embodiment of the invention the type III effector moiety is linked to the translocation domain via a disulphide bridge.

In a preferred embodiment of the invention, the targeting domain is derived from a clostridial neurotoxin binding fragment (H_c domain). This may be derived from tetanus toxin or any one of the eight botulinum toxin serotypes (A-G). Delivery via the clostridial neurotoxin receptors differs significantly to the normal delivery route of the type III effectors and offers a number of advantages:

The clostridial H_c fragments bind with high affinity to receptors on the cell surface and provide high specificity to neuronal cells. The clostridial neurotoxins are internalised via an acidic endosome which triggers the translocation of the type III effector moiety across the membrane and into the cytosol.

For non-neuronal cells a wide range of high affinity binding domains have been defined for specific cell types. Examples are described for a number of cellular targets.

The agent can comprise a ligand or targeting domain, which binds to an endocrine cell and is thus rendered specific for these cell types. Examples of suitable ligands include iodine; thyroid stimulating hormone (TSH); TSH receptor antibodies; antibodies to the islet-specific monosialo-ganglioside GM2-1; insulin, insulin-like growth factor and antibodies to the receptors of both; TSH releasing hormone (protirelin) and antibodies to its receptor; FSH/LH releasing hormone (gonadorelin) and antibodies to its receptor; corticotrophin releasing hormone (CRH) and antibodies to its receptor; and ACTH and antibodies to its receptor.

Ligands suitable to target an agent to inflammatory cells include (i) for mast cells, complement receptors in general, including C4 domain of the Fc IgE, and antibodies/ligands to the C3a/C4a-R complement receptor; (ii) for eosinophils,

antibodies/ligands to the C3a/C4a-R complement receptor, anti VLA-4 monoclonal antibody, anti-IL5 receptor, antigens or antibodies reactive toward CR4 complement receptor; (iii) for macrophages and monocytes, macrophage stimulating factor, (iv) for macrophages, monocytes and neutrophils, bacterial LPS and yeast B-glucans which bind to CR3, (v) for neutrophils, antibody to OX42, an antigen associated with the iC3b complement receptor, or IL8; (vi) for fibroblasts, mannose 6-phosphate/insulin-like growth factor-beta (M6P/IGF-II) receptor and PA2.26, antibody to a cell-surface receptor for active fibroblasts in mice.

Ligands suitable to target an agent to exocrine cells include pituitary adenyl cyclase activating peptide (PACAP-38) or an antibody to its receptor.

Ligands suitable to target an agent to immunological cells include Epstein Barr virus fragment/surface feature or idiotypic antibody (binds to CR2 receptor on B-lymphocytes and lymph node follicular dendritic cells).

Suitable ligands for targeting platelets for the treatment of disease states involving inappropriate platelet activation and thrombus formation include thrombin and TRAP (thrombin receptor agonist peptide) or antibodies to CD31/PECAM-1, CD24 or CD106/VCAM-1, and ligands for targeting cardiovascular endothelial cells for the treatment of hypertension include GP1b surface antigen recognising antibodies.

Suitable ligands for targeting osteoblasts for the treatment of a disease selected from osteopetrosis and osteoporosis include calcitonin, and for targeting an agent to osteoclasts include osteoclast differentiation factors (eg. TRANCE, or RANKL or OPGL), and an antibody to the receptor RANK.

In one embodiment of the invention the translocation domain is derived from a bacterial toxin. Examples of suitable translocation domains are those derived from the clostridial neurotoxins or diphtheria toxin.

In another embodiment of the invention, the translocation domain is a membrane disrupting or 'fusogenic' peptide, which functions as a translocation domain. An example of such a peptide is that derived from influenza virus haemagglutinin HA2 (residues 1-23).

In one example of the construct of the invention, the type III effector protein is SigD from *Salmonella spp.* In another example of the construct of the invention, the type III effector protein is YopE from *Yersinia spp.*

In an example of the construct of the invention in which the type III effector moiety is SigD from *Salmonella spp*, the construct may consist of the following:-

the SigD type III effector moiety;

the translocation domain from diphtheria toxin;

the binding domain (H_c domain) from botulinum type A neurotoxin; and

a linker peptide to enable attachment of the SigD effector to the translocation domain via a disulphide bridge.

In an another example of the construct of the invention in which the type III effector moiety is SigD from *Salmonella spp,* the construct consists of the following:-

the SigD type III effector moiety;

the translocation domain in the form of a fusogenic peptide;

the binding domain (H_c domain) from botulinum type F neurotoxin; and

a linker peptide to enable attachment of the SigD effector to the translocation domain via a disulphide bridge.

In an example of the construct of the invention in which the type III effector moiety is YopE from *Yersinia spp*, the construct may consist of the following:-

the YopE type III effector moiety;

the translocation domain from diphtheria toxin;

the binding domain (H_c domain) from botulinum type F neurotoxin; and

a linker peptide to enable attachment of the YopE effector to the translocation domain via a disulphide bridge.

The invention enables manipulation of cell signalling, and in a specific example SigD is incorporated into a construct of the invention and can be used to promote neuronal cell survival under stress. By targeting the appropriate intracellular signalling pathway, it is possible to simultaneously regulate a number of pathways to improve the prospects for neuronal survival. SigD (also known as SopB) activates the protein kinase Akt, which is a key intermediate in the pro-survival signalling pathways mediated by various growth factors. Not only does Akt up-regulate pro-survival transcription factors such as NF-κB,

but it also down-regulates several pro-apoptotic factors such as Bad and Forkhead.

A number of type III and IV effectors function to maintain the intracellular niche of the bacteria within the host cell. Whilst some bacterial pathogens are released into the cell cytosol, many form and maintain a specialised intracellular compartment sometimes termed a vacuole. One of the principle functions of many effector protein is to regulate the fusion of the compartment with other intracellular compartments such as potentially damaging phagolysosomal. At the same time the pathogen may need to promote fusion with other membrane bound compartments, including recycling endosomes, to either provide nutrients to the encapsulated pathogen or allow the dissemination of the pathogen to other locations. Intracellular pathogens offer a wide range of effector molecules for regulating intracellular trafficking and membrane fusion.

The mechanism underlying the fusion of membrane bound vesicles is conserved in a number of cellular processes. Broadly speaking, membrane fusion events are classified either as secretory processes for the release of material from the plasma membrane, or as endocytic processes that move material from the plasma membrane to the lysosomal system. This simplified classification does not take into account retrograde and anterograde processes, which occur within these pathways, or multiple points of communication between the two pathways. The underlying mechanism in all membrane fusion events can be broken down into 4 component phases:

The transported material is concentrated at a specific site on the donor membrane and is "pinched off" in a vesicle that becomes separated from this membrane.

The vesicle is transported to the acceptor membrane along cytoskeletal fibres (e.g. microtubules).

The vesicle then attaches to the acceptor membrane via a "docking/tethering" mechanism mediated by SNARE complex proteins.

The vesicle and the acceptor membrane fuse to release the contents of the

vesicle through the acceptor membrane.

Thus similar SNARE proteins and regulatory proteins underpin the fusion of endosomal vesicles with the lysosome, endoplasmic reticulum with the Golgi and trans-Golgi network, and secretory vesicles with the plasma membrane. The functional conservation of the membrane fusion mechanism means that a bacterial effector protein that would normally regulate the fusion of a specific event can be directed to modulate other fusion events. For example, an effector that blocks endosomal fusion with the lysosome can be redirected to block the fusion of secretory vesicles with the plasma membrane, or ER vesicles with the Golgi network.

One of the key classes of regulatory proteins that have been defined in vesicle trafficking are small GTPases of the Ras superfamily termed Rab proteins (or Ypt proteins in yeast). Rab proteins are implicated in every stage of membrane fusion. For example Rab 1,2,5 and 9 are involved in sorting material for transport (stage 1 above), Rab5,6,27 and Sec4 mediate transport (stage 2), Rab1,5, Ypt1,7 Sec4 influence docking to the acceptor membrane (stage 3) and other Rab proteins implicated in promoting membrane fusion. The list above shows that certain Rab proteins, such as Rab1 and Rab5, are involved in more than 1 stage of the fusion process. Similarly some Rab proteins are present on all membrane vesicles whilst others have more specialised roles in specific fusion events.

Rab proteins are key potential targets for modification by either bacterial pathogens intent on blocking or promoting membrane fusion events or by therapeutic agents designed to regulate intracellular trafficking. One of the first effector proteins to be described as having an effect on Rab function was the secreted effector protein SopE2 from *Salmonella* species. SopE2 acts as a guanine nucleotide exchange factor for Rab5a resulting in increased activation of the protein on the cell membrane. This activity has been correlated with increased survival of Salmonella in infected HeLa cells and macrophages (Cell Micobiol. 3 p473). SpiC is another Salmonella effector that blocks endosome fusion (EMBO J. 18p3924-3933). Unlike SopE, which shows some conservation with normal cellular regulators of GTPase, SpiC shows no clear homology to other proteins. Its ability to block one of the four stages of vesicle fusion is known. It could exert its activity at the level of the SNARE proteins,

modulate Rab function directly or operate at the level of one of the regulators of Rab function. Membrane insertion is essential for Rab activity. Rab proteins form a stable complex with Rab escort protein (REP) in the cytosol and this is a substrate for a geranyl geranyl transferase (RabGGT) which adds a C-terminal isoprenoid moiety. In the absence of REP or RabGGT the Rab protein would remain in an inactive form in the cytosol. REP also mediates the membrane insertion of the modified Rab into the donor membrane. Rab proteins can also be retrieved from the membrane via the action of Rab GDP dissociation inhibitor (RabGDI). All of these proteins are potential targets for bacterial pathogens to alter membrane fusion events. The precise effect would depend on whether alterations cause an increase or decrease in the levels of active Rab in the donor membrane, and the specificity for particular Rab proteins.

A number of human diseases have now been identified in which mutations affect either Rab proteins or their regulators. These human diseases serve to illustrate the cellular effects of alterations in Rab control in cells. Thus mutations in Rab27 (Griscelli syndrome), REP1 (choroiderma), RabGDIα (X-linked mental retardation) and RabGGT α subunit (Hermansky-Pudlack syndrome) are all implicated in human disease (as reviewed in Seabra *et al Trends in Molecular Medicine* (2002) **8**;23-26, Olkkonen and Ikonen *New England Journal of Medicine* (2000) **343**;1104)). A wide range of human diseases involve defects in intracellular trafficking (as reviewed in Aridor and Hannan *Traffic* (2000) **1**;836-851). Modulation of membrane fusion via the specialised properties of bacterial effector proteins directed at one of the 4 mechanisms described above offers therapeutic opportunities for these diseases and others where transport properties are affected.

The targeting of the membrane fusion event between secretory vesicles and the plasma membrane allow the control of secretion from cells. Effectors that alter regulation of specific Rab proteins, either directly or via one of the mechanisms described above, including Rab3a,b,c and d, Rab8a and b, Rab26, Rab27a Rab37, or affect any of the other molecular events of membrane fusion (1-4 described above) can regulate secretion. Effector proteins can be applied to either increase or decrease secretion from a specific cell type. In a therapeutic context this is valuable for the treatment of a wide range of disorders including muscle spasms (blephorospasm, torticolis etc)

hypersecretion disorders (COPD, bronchitis, asthma).

By modulating the fusion of recycling endosomes with either the lysosome or the plasma membrane it also possible to modulate the presentation of specific families of cell surface marker. Again effectors directed to alter regulation of specific Rab proteins, such as Rab4a and b, Rab11a and b, Rab15, Rab17, Rab18 or affect other molecular events in the fusion mechanism, can either up or down regulate presentation of cell surface marker. Therapeutically this has enormous potential for altering the response of cells to external stimuli (e.g. modulating response to growth factors, hormones, cytokines, chemokines or other signalling molecules), modifying the recognition of cells by external factors (e.g. immune surveillance) or for switching cell signalling pathways on or off.

Using constructs of the invention, therapeutic intervention can be provided in neurodegenerative disorders such as Alzheimer's disease and Prion diseases (vCJD). Both diseases are characterised by the accumulation of insoluble protein plaques due to misfolding of cellular proteins. In both cases an intracellular amplification of misfolded protein, via passage through endosomal-lysosomal compartments, is implicated in the progression of the disease. Neuronally targeted bacterial effectors as described herein, which modulate the fusion of endosomal and lysosomal compartments, allow control of the accumulation of insoluble protein. As this is one of the key survival strategies of many intracelullar bacterial pathogens, a number of therapeutic molecules are available, for example Salmonella effectors such as SpiC, SptP and SopE2.

In still further embodiments of the invention, constructs are provided for inhibition or promotion of secretion, containing a type III effector and a targetting moiety. Specific effectors for this purpose are selected from SpiC, SopE, RalF, Sse E, F, G and J, PipA, PipB, SifA and SifB. These constructs target the membrane fusion event between secretory vesicles and the plasma membrane to allow the control of secretion from cells. Effectors that alter regulation of specific Rab proteins, either directly or via one of the mechanisms described above, including Rab3a,b,c and d, Rab8a and b, Rab26, Rab27a Rab37, or affect any of the other molecular events of membrane fusion, can regulate secretion. Effector proteins can be applied to either increase or

decrease secretion from a specific cell type. In a therapeutic context this is valuable for the treatment of a wide range of disorders including muscle spasms (blephorospasm, torticolis etc) hypersecretion disorders (COPD, bronchitis, asthma).

The pathogenic strategy to establish a specialised intracellular niche and to modulate fusion of that compartment with other vesicles is conserved for a vast range of pathogens. Not only does this provide a vast range of molecules capable of modulating the cellular events as described above, but it also provides an array of potential therapeutic targets for such molecules. Although many of the intracellular pathogens described in table 2 establish membrane bound compartments, the precise biochemistry and the signalling events and effectors needed to maintain these compartments are very different. A few intracellular pathogens escape from the phagosomal or endosomal compartment in which they enter the cell. The effector proteins involved in this process are incompatible with the life cycle of pathogens that remain in membrane compartments. The effector proteins of two intracellular pathogens existing in membrane bound vesicles are also not necessarily compatible. For example, enhancement of Rab5a activity by Salmonella in macrophages is correlated with enhanced survival (Cell Microbiology **3**;473-). However, increases in Rab5a expression/activity accelerates intracellular destruction of *Listeria monocytogenes* in macrophages (J. Biological Chemistry **274**;11459). The Salmonella effector proteins that are likely to be involved in Rab5a recruitment (e.g. SopE2, SpiC or other SPI-2 secreted effectors) are therefore potential therapeutic agents for treating intracellular *Listeria*.

In its crudest form anti-microbial therapy could involve treating one intracellular pathogen with a second pathogen on the basis that the two intracellular compartments and requirements of the organisms would not be compatible. For example treatment of TB infected macrophages with Salmonella might be expected to result in provoked "vacuole" lysosome fusion within the macrophage leading to the eradication of the TB. The type and fate of the super-infecting pathogen would have to be carefully chosen so as not to exacerbate the infectivity or spread of the original organism.

A refinement of the superinfection strategy would therefore focus on the targeted delivery of effector molecules to specific target cells as described by

this invention. This could either utilise a highly attenuated pathogen (e.g. Salmonella that only secretes SopE2 or SptP) or targeted protein delivery (e.g. using a toxin delivery domain, antibody or similar cell targeting ligands). Protective antigen from *Bacillus anthracis* would be capable of targeting effectors to macrophages for the treatment of a wide range of bacterial pathogens. The specific addition of carbohydrate moieties will enable specific targeting of pools of macrophages via the mannose receptor (e.g Vyas et al, International Journal of Pharmaceutics (2000) **210**p1-14). A cell surface marker specific for infected cells (as distinct from uninfected cells) would offer an ideal target for delivery systems. The cell type infected by the pathogen would determine the choice of delivery ligand whilst the precise fate of the cell compartment would determine the choice of effector (e.g. cell apoptosis, lysis, endosome-lysosome fusion, endosome acidification etc).

A key benefit of this type of therapy is that the effector proteins are not intrinsically toxic to the cell and therefore delivery of the protein to uninfected target cells is unlikely to have any deleterious effects. In this case, whilst desirable, the precise specificity of the targeting ligand is not essential for successful therapy.

The wide range of intracellular pathogens and the difficulty in treating/immunising against these organisms make this approach a valuable alternative to antibiotic therapy. The method is also attractive as avoidance of the antimicrobial agent either means that the pathogen must produce a molecule capable of overriding the effector-induced cell stimulus or must significantly modify its lifestyle. For obligate intracellular pathogens or where the intracellular stage is essential for propagation, this may offer greater hopes for extended antimicrobial use than current antibiotic strategies targeted at specific biochemical interactions.

In another example of the invention in which the effector protein is SpiC from *Salmonella* spp, the construct may consist of the following:-

the SpiC effector moiety fused to a domain capable of interacting with protective antigen;

the protective antigen from *Bacillus anthracis*;

where the construct is either co-administered or where the SpiC moiety is administered after the protective antigen.

The constructs of the invention are preferably produced either wholly or partially by recombinant technology. In an embodiment of the invention where a construct of the invention is produced by recombinant technology, the construct of the invention will be produced as a single multi-domain polypeptide comprising from the N-terminus:-

the type III effector moiety;

a linker peptide;

the translocation domain; and

the binding domain.

In such a construct, the C-terminus of the type III effector protein is fused to the N-terminus of the translocation domain via the linker peptide. An example of such a linker peptide is the sequence CGLVPAGSGP which contains the thrombin protease cleavage site and a cysteine residue for disulphide bridge formation. The latter single chain fusion protein may then be treated with thrombin to give a dichain protein in which the type III effector is linked to the translocation domain by a disulphide link. In another example of a linker peptide in which the translocation domain does not contain a free cysteine residue near its C-terminus, such as is the case when the translocation domain is a fusogenic peptide, the linker peptide contains both cysteine residues required for the disulphide bridge. An example of the latter linker peptide is the amino acid sequence: CGLVPAGSGPSAGSSAC.

In an example of the construct of the invention in which the type III effector moiety is SigD from *Salmonella spp* produced by recombinant technology, the construct may consist of polypeptide containing (from the N-terminus) the following domains:-

the SigD type III effector moiety;

linker peptide (sequence CGLVPAGSGP) to enable attachment of the SigD effector to the translocation domain via a disulphide bridge;

the translocation domain from diphtheria toxin (residues 194-386); and

the binding domain (H_c domain) from botulinum type A neurotoxin (residues 872-1296).

The constructs of the invention may also be produced using chemical cross-linking methods. Various strategies are known by which type III effector

proteins can be linked to a polypeptide consisting of the translocation domain and binding domain using a variety of established chemical cross-linking techniques. Using these techniques a variety of type III effector constructs can be produced. The type III effector protein is, for example, derivatised with the cross-linking reagent N-succinimidyl 3-[2-pyridyldithio] propionate. The derivatised type III effector is then linked to a peptide containing a translocation domain and binding domain via a free cysteine residue present on the translocation domain.

Protein effectors can be altered to allow their delivery to intracellular compartments other than their usual site of action. For example, mitochondrial or nuclear targeting signals could be added to direct the effector to these compartments. By covalently linking the effector to the targeting domain the effector can be retained in the endosome/lysosome compartment, which would not normally be accessible by bacterial delivery. Effectors can be targeted to specific membrane locations via lipid modifications including myristoylation, palmitoylation, or the addition of short proteins domains that might include SH2, SH3, WW domains, fragments of Rab proteins or synaptojanin-like domains. Those practised in the art would recognise that these targeting strategies offer an advantage for certain therapeutic strategies.

Constructs of the invention may be introduced into either neuronal or non-neuronal tissue using methods known in the art. By subsequent specific binding to neuronal cell tissue, the targeted construct exerts its therapeutic effects. Ideally, the construct is injected near a site requiring therapeutic intervention.

The construct of the invention may be produced as a suspension, emulsion, solution or as a freeze dried powder depending on the application and properties of the therapeutic substance. The construct of the invention may be resuspended or diluted in a variety of pharmaceutically acceptable liquids depending on the application.

"Clostridial neurotoxin" means either tetanus neurotoxin or one of the seven botulinum neurotoxins, the latter being designated as serotypes A, B C_1, D, E, F or G, and reference to the domain of a toxin is intended as a reference to the intact domain or to a fragment or derivative thereof which retains the essential function of the domain.

"Conjugate" means, in relation to two polypeptides, that the polypeptides are linked by a covalent bond, typically forming a single polypeptide as a result, or by a di-sulphide bond.

"Binding domain" means a polypeptide which displays high affinity binding specific to a target cell, e.g. neuronal cell binding corresponding to that of a clostridial neurotoxin. Examples of binding domains derived from clostridial neurotoxins are as follows:-

Botulinum type A neurotoxin	– amino acid residues (872 - 1296)
Botulinum type B neurotoxin	– amino acid residues (859 - 1291)
Botulinum type C neurotoxin	– amino acid residues (867 - 1291)
Botulinum type D neurotoxin	– amino acid residues (863 - 1276)
Botulinum type E neurotoxin	– amino acid residues (846 - 1252)
Botulinum type F neurotoxin	– amino acid residues (865 - 1278)
Botulinum type G neurotoxin	– amino acid residues (864 - 1297)
Tetanus neurotoxin	– amino acid residues (880 - 1315)

"High affinity binding specific to neuronal cell corresponding to that of a clostridial neurotoxin" refers to the ability of a ligand to bind strongly to cell surface receptors of neuronal cells that are involved in specific binding of a given neurotoxin. The capacity of a given ligand to bind strongly to these cell surface receptors may be assessed using conventional competitive binding assays. In such assays radiolabelled clostridial neurotoxin is contacted with neuronal cells in the presence of various concentrations of non-radiolabelled ligands. The ligand mixture is incubated with the cells, at low temperature (0-3°C) to prevent ligand internalization, during which competition between the radiolabelled clostridial neurotoxin and non-labelled ligand may occur. In such assays when the unlabelled ligand used is the same as that of the labelled neurotoxin, the radiolabelled clostridial neurotoxin will be displaced from the neuronal cell receptors as the concentration of non-labelled neurotoxin is increased. The competition curve obtained in this case will therefore be representative of the behaviour of a ligand which shows "high affinity binding specificity to neuronal cells corresponding to that of a clostridial neurotoxin", as used herein.

A carrier that "targets" a particular cell generally does so due to binding of the

carrier, or a portion thereof, to that cell and, by way of example, many different ligands with given cell type specificity are described herein.

"Translocation domain" means a domain or fragment of a protein which effects transport of itself and/or other proteins and substances across a membrane or lipid bilayer. The latter membrane may be that of an endosome where translocation will occur during the process of receptor-mediated endocytosis. Translocation domains can frequently be identified by the property of being able to form measurable pores in lipid membranes at low pH (Shone *et al.* Eur J. Biochem. 167, 175-180). Examples of translocation domains are set out in more detail below:

Diphtheria toxin	– amino acid residues (194 - 386)
Botulinum type A neurotoxin	– amino acid residues (449 - 871)
Botulinum type B neurotoxin	– amino acid residues (441 - 858)
Botulinum type C neurotoxin	– amino acid residues (442 - 866)
Botulinum type D neurotoxin	– amino acid residues (446 - 862)
Botulinum type E neurotoxin	– amino acid residues (423 - 845)
Botulinum type F neurotoxin	– amino acid residues (440 - 864)
Botulinum type G neurotoxin	– amino acid residues (442 - 863)
Tetanus neurotoxin	– amino acid residues (458 - 879)

Translocation domains are frequently referred to herein as "H_N domains".

"Translocation" in relation to translocation domain, means the internalization events that occur after binding to the cell surface. These events lead to the transport of substances into the cytosol of target cells.

"Injected effector secreted by type III or type IV secretion system" means bacterial proteins that are injected into host cells (mammalian, plant, insect, fish or other) via a modified pilus or "needle-like" injection system frequently referred to as type III or type IV secretion systems" and the term embraces fragments, modifications and variations thereof that retain the essential effector activity.

The invention thus uses modification of intracellular signalling for promoting neuronal growth. Many of the effectors and inhibitors that control the

development of the growth cone act through common intracellular signalling pathways that modulate the phosphorylation state of cytoskeletal components and that ultimately determine whether the axon grows or collapses. The appropriate manipulation of intracellular signalling is therefore a powerful approach for eliminating the need for multiple inhibitors of the many factors shown to induce apoptosis and growth cone collapse. The up-regulation of transcription factors that inhibit apoptosis is an example of manipulation of intracellular signalling to promote neural regeneration.

Strategies for therapeutic intervention using the effectors and compositions of the invention include the delivery of agents to eliminate stress-inducing factors and the modification of intracellular signalling to promote cell survival. The latter approach is particularly powerful and the present invention describes conjugates with type III effector moieties which allow such strategies to be pursued.

The constructs of this invention use a specific targeting system to ensure delivery of the therapeutic agent to the desired cells and uses bacterial toxins that have evolved to regulate key stages in the cell signalling machinery of the cells. This strategy offers a number of advantages over other drug platforms. The cell specificity ensures that any alterations in cell signalling occur only in the cells where this modification is desirable and not in other adjacent cells. For example, in neuronal cell-targeted constructs, changes in signalling would only take place in neurones and not in adjacent glial cells where such changes might not be desirable. By targeting key intermediates in the signalling pathway it is possible to co-ordinately regulate a number of overlapping cellular events to promote the desired effect. For example, the activation of Akt by SigD causes an effect on cells by co-ordinating a number of signalling pathways to actively promote cell survival and block the induction of apoptosis in response to stress factors. This is also a good example of an effector that activates a component of a cell-signalling pathway. Most drug discovery approaches tend to identify inhibitors of specific components.

The invention is now illustrated in the following specific examples.

Examples:

Example 1. Cloning and expression of type III effector genes.

Standard molecular biology protocols were used for all genetic manipulations (Sambrook *et al* 1989, Molecular cloning; A laboratory manual. Second Edition, Cold Spring Harbor Laboratory Press, New York.). Genes encoding Type III effectors, truncated versions removing the N-terminal hydrophobic domain (e.g removal of amino acids 1-28 for SigD, 1-69 for SptP, 1-76 for SopE), or individual sub-domains (e.g. ExoS GAP domain amino acids 96-234 and ADP-ribosyltransferase domain amino acids 232-453), were amplified from genomic DNA by PCR to generate suitable restriction sites for cloning. In some cases synthetic genes were prepared with codon usage optimised for expression in *E.coli*. Restriction enzymes such as *Bam*HI (5') and *Bg*lI (3') were used for cloning with reading frames maintained. Constructs were sequenced to confirm the presence of the correct sequence. Constructs for expression were subcloned, as a suitable fragment, into an expression vector carrying a T7 polymerase promoter site (e.g. pET28, pET30 or derivatives (Novagen Inc, Madison, WI)), to generate a fusion with maltose binding protein (e.g. pMALc2x (NEB)) or into other expression vectors known to those familiar with the art. Clones with confirmed sequences were used to transform expression hosts: For T7 polymerase vectors *E.coli* BL21 (DE3) (Studier and Moffatt 1986 *Journal of Molecular Biology* 189:113-130) JM109 (DE3) or equivalent strains with a DE3 lysogen. For pMalc2x JM109, BL21, TG1, TB1 or other suitable expression strains.

In addition to the expression of type III effectors as standard fusion proteins an additional approach was used to generate fusion proteins. The type III effector or truncated effector generated as above were cloned into the 5' end of a gene encoding a cell targeting ligand, which include toxin fragments, antibodies, growth factors, lectins, interleukins, peptides. These fusion proteins were cloned and expressed as either 6-His tagged, MBP tagged or other fusions as described above.

Expression cultures were grown in Terrific Broth containing 30µg/ml kanamycin and 0.5% (w/v) glucose to an OD_{600} of 2.0 and protein expression was induced with 500µM IPTG for 2 hours. Cells were lysed by either sonication or suitable detergent treatment (e.g. Bugbuster reagent; Novagen), cell debris pelleted by centrifugation and the supernatant loaded onto a metal chelate column charged

with Cu^{2+} (Amersham-Pharmacia Biotech, Uppsala, Sweden).

The recombinant proteins expressed from pET vectors contain amino-terminal histidine (6-His) and T7 peptide tags allowing proteins to be purified by affinity chromatography on either a Cu^{2+} charged metal chelate column. After loading proteins on the column and washing, proteins were eluted using imidazole. All buffers were used as specified by manufacturers. Where appropriate removal of the purification tag was carried out according to manufacturers instructions.

MBP fusions were purified on amylose resin columns as described by the manufacturer (NEB) following growth in Terrific Broth containing 100 µg/ml ampicillin and lysis as described above.

Other fusion systems were used according to manufacturer's instructions and purification carried out on suitable columns using defined methods.

Example 2. Production of recombinant targeting vectors consisting of translocation and binding domains

Standard molecular biology protocols were used for all genetic manipulations (Sambrook *et al* 1989, Molecular cloning; A laboratory manual. Second Edition, Cold Spring Harbor Laboratory Press, New York.) Clostridial neurotoxin binding domains (BoNT/Hc or TeNT/Hc) derived from either their native genes or synthetic genes with codon usage optimised for expression in *E.coli* were amplified by PCR. Introduced *Bam*HI (5') restriction sites and *Hind*III, *Sal*I or *Eco*RI (3') sites were used for most cloning operations with reading frames designed to start with the first base of the restriction site. Constructs were sequenced to confirm the presence of the correct sequence. The translocation domain of diphtheria toxin (DipT) was amplified from its native gene to introduce *Bam*HI and *Bgl*II sites at the 5' and 3' ends respectively. This *Bam*HI and *Bgl*II fragment was subcloned into the *Bam*HI site at the 5' end of the Hc fragment to generate an in-frame fusion. The entire heavy chain fragment (DipT-Hc) was excised as a *Bam*HI-*Hind*III or *Bam*HI-*Sal*I or *Bam*HI-*Eco*RI fragment and subcloned into a suitable expression vector.

Constructs for expression were subcloned into either an expression vector carrying a T7 polymerase promoter site (e.g. pET28, pET30 or derivatives

(Novagen Inc, Madison, WI)) or to generate a fusion with maltose binding protein (e.g. pMALc2x (NEB)) as a suitable fragment. Clones with confirmed sequences were used to transform expression hosts: For T7 polymerase vectors *E.coli* BL21 (DE3) (Studier and Moffatt 1986 *Journal of Molecular Biology* 189:113-130) JM109 (DE3) or equivalent strains with a DE3 lysogen. For pMalc2x JM109, BL21, TG1, TB1 or other suitable expression strains.

The recombinant proteins expressed from pET vectors contain amino-terminal histidine (6-His) and T7 peptide tags allowing proteins to be purified by affinity chromatography on either a Cu^{2+} charged metal chelate column. Expression cultures were grown in Terrific Broth containing 30microg/ml kanamycin and 0.5% (w/v) glucose to an OD_{600} of 2.0 and protein expression was induced with 500microM IPTG for 2 hours. Cells were lysed by either sonication or suitable detergent treatment (e.g. Bugbuster reagent; Novagen), cell debris pelleted by centrifugation and the supernatant loaded onto a metal chelate column charged with Cu^{2+} (Amersham-Pharmacia Biotech, Uppsala, Sweden). After loading proteins on the column and washing, proteins were eluted using imidazole. All buffers were used as specified by manufacturers. Where appropriate removal of the purification tag was carried out according to manufacturers instructions.

MBP fusions were purified on amylose resin columns as described by the manufacturer (NEB) following growth in Terrific Broth containing 100 microg/ml ampicillin and lysis as described above.

Thrombin or factor Xa protease sites were included within the protein for subsequent removal of these purification tags.

Additional sequences for adding affinity purification tags and one or more specific protease sites for the subsequent removal of these affinity tags may also be included in the reading frame of the gene products.
Other coding sequences that enable expression of the desired protein would also be acceptable. Other tags or linking sites may also be incorporated into the sequence.

Using the techniques described above, targeting vector fragments were constructed by fusing domains of the H_c fragments of either botulinum type A, type F or tetanus neurotoxins with the translocation domain of diphtheria toxin.

Example 3. Preparation of botulinum heavy chains by chemical methods.

The various serotypes of the clostridial neurotoxins may be prepared and purified from various toxigenic strains of *Clostridium botulinum* and *Clostridium tetani* by methods employing standard protein purification techniques as described previously (Shone and Tranter 1995, Current Topics in Microbiology, 194, 143-160; Springer). Samples of botulinum neurotoxin (1mg/ml) are dialysed against a buffer containing 50mM Tris-HCl pH 8.0, 1M NaCl and 2.5M urea for at least 4 hours at 4°C and then made 100mM with dithiothreitol and incubated for 16h at 22°C. The cloudy solution, which contains precipitated light chain, is then centrifuged at 15000 x g for 2 minutes and the supernatant fluid containing the heavy chain retained and dialysed against 50mM HEPES pH 7.5 containing 0.2M NaCl and 5mM dithiothreitol for at least 4 hours at 4°C. The dialysed heavy chain is centrifuged at 15000 x g for 2 minutes and the supernatant retained and dialysed thoroughly against 50mM HEPES pH 7.5 buffer containing 0.2M NaCl and stored at -70°C. The latter procedure yields heavy chain >95% pure with a free cysteine residue which can be used for chemical coupling purposes. Biological (binding) activity of the heavy chain may be assayed as described in Example 5.

The heavy chains of the botulinum neurotoxins may also be produced by chromatography on QAE Sephadex as described by the methods in Shone and Tranter (1995) (Current Topics in Microbiology, 194, 143-160; Springer).

Example 4. Chemical conjugation of proteins

Recombinant SigD type III effector from *Salmonella spp.* was cloned and purified as described in Example 1. The SigD type III effector was chemically modified by treatment with a 3-5 molar excess of N-succinimidyl 3-[2-pyridyldithio] propionate (SPDP) in 0.05M Hepes buffer pH 7.0 containing 0.1M NaCl for 60 min at 22°C. The excess SPDP was removed by dialysis against the same buffer at 4°C for 16h. The substituted SigD effector was then mixed in a 1:1 ratio and incubated at 4°C for 16h with a targeting vector comprising a translocation domain (with an available free cysteine residue) and a neuronal targeting domain (see Example 2). The latter may also be native heavy chain purified from *Clostridium botulinum* type A neurotoxin purified as described in

Example 3. During the incubation period the SigD effector was conjugated to the targeting vector fragment by a free sulphydryl group. After incubation, the SigD-construct was purified by gel filtration chromatography on Sephadex G200.

Example 5. Assay of the biological activity of constructs - demonstration of high affinity binding to neuronal cells.

Clostridial neurotoxins may be labelled with 125-iodine using chloramine-T and its binding to various cells assessed by standard methods such as described in Evans *et al.* 1986, Eur J. Biochem., 154, 409 or Wadsworth *et al.* 1990, Biochem. J. 268, 123). In these experiments the ability of Type III constructs to compete with native clostridial neurotoxins for receptors present on neuronal cells or brain synaptosomes was assessed. All binding experiments were carried out in binding buffers. For the botulinum neurotoxins this buffer consisted of: 50mM HEPES pH 7.0, 30mM NaCl, 0.25% sucrose, 0.25% bovine serum albumin. For tetanus toxin, the binding buffer was: 0.05M tris-acetate pH 6.0 containing 0.6% bovine serum albumin. In a typical binding experiment the radiolabelled clostridial neurotoxin was held at a fixed concentration of between 1-20nM. Reaction mixtures were prepared by mixing the radiolabelled toxin with various concentrations of unlabelled neurotoxin or construct. The reaction mixtures were then added to neuronal cells or rat brain synaptosomes and then incubated at 0-3°C for 2hr. After this period the neuronal cells of synaptosomes were washed twice with binding ice-cold binding buffer and the amount of labelled clostridial neurotoxin bound to cells or synaptosomes was assessed by â-counting. In an experiment using an Type III effector construct what contained the binding domain from botulinum type A neurotoxin, the construct was found to compete with ^{125}I-labelled botulinum type A neurotoxin for neuronal cell receptors in a similar manner to unlabelled native botulinum type A neurotoxin. These data showed that the construct had retained binding properties of the native neurotoxin.

Example 6. Recombinant Type III effector constructs

Recombinant Type III effector-targeting vector constructs were prepared comprising a combination of the following elements:-

- a type III effector (e.g. SigD from *Salmonella spp.*)

- a linker region, which allows the formation of a disulphide bond between the type III effectors and the translocation domain and which also contains a unique protease cleavage site for cleavage by factor Xa or thrombin to allow the formation of a dichain molecule;

- a translocation domain from diphtheria toxin or a endosomolytic (fusogenic) peptide from influenza virus haemagglutinin); and

- a neuronal cell-specific binding domain (e.g. from tetanus or botulinum neurotoxin type A or botulinum neurotoxin type F).

The protein sequences of these various domains form specific embodiments of the invention and are shown below the examples.

To confirm the nature of their structure, the recombinant Type III effector-targeting vector constructs were converted to the dichain form by treatment with a unique protease corresponding to the cleavage site sequences within the linker region. Conjugates containing the thrombin cleavage site were treated with thrombin (20microg per mg of conjugate) for 20h at 37°C; conjugates containing the factor Xa cleavage site were treated with factor Xa (20microg per mg of conjugate) for 20 min at 22°C.

On SDS-PAGE gels, under non-reducing conditions, the majority of Type III effector-targeting vector construct appeared as single band. In the presence of reducing agent (dithiothreitol) two bands were observed corresponding to the type III effector and targeting vector (translocation and binding domains). These data illustrate that, after treatment with the unique protease, the conjugates consist of the latter two components which are linked by a disulphide bridge.

Example 7. Formation of Type III effector constructs from Type III effector-diphtheria toxin A (CRM197) fusion proteins.

Type III effector-targeting vector constructs may also formed *in vitro* by reconstitution from two recombinant fragments. These are:-

A Type III effector fused to inactive diphtheria fragment A (CRM197) as described in Example 1.

A recombinant targeting vector in which the translocation domain of diphtheria toxin is fused to a neuronal targeting domain such as that from a clostridial neurotoxin. Production of these is described in Example 2.

Type III effector constructs may be formed by mixing fragments 1 and 2 in equimolar proportions in the presence of 10mM dithiothreitol and them completely removing the reducing agent by dialysis against phosphate buffered saline at pH 7.4 followed by dialysis against HEPES (0.05M, pH 7.4) containing 0.15 M NaCl. As described above in Example 6, these constructs appear as a single band in SDS gels under non-reducing conditions and two bands in the presence of a reducing agent.

Example 8. Formulation of the Type III effector construct for clinical use.

In a formulation of the Type III effector construct for clinical use, recombinant Type III effector construct would be prepared under current Good Manufacturing Procedures. The construct would be transferred, by dilution, to a solution to give the product stability during freeze-drying. Such a formulation may contain Type III effector construct (concentration between 0.1 -10 mg/ml) in 5mM HEPES buffer (pH 7.2), 50mM NaCl, 1% lactose. The solution, after sterile filtration, would be aliquotted, freeze-dried and stored under nitrogen at -20°C.

Example 9. Production of constructs with neuroprotective properties.

SigD was cloned (without the first 29 condons) using the methods outlined in Example 1. The protein was expressed and purified either as a fusion with maltose binding protein (e.g. using pMALc2x) or with a Histidine6 (e.g. using pET28a). Purification tags were then removed by standard procedures after affinity purication of the fusion protein. Chemical constructs of SigD were prepared as outlined in Example 4.

A recombinant construct of the invention containing SigD linked to the translocation domain and binding domain of botulinum type A neurotoxin was prepared as outlined in Example 6 using the standard molecular biology procedures outlined in Example 1.

Application of the above constructs to neuronal cells leads to the receptor-mediated internalisation of SigD and subsequent activation of Akt Kinase. Such cells have an enhanced ability to withstand stress such as growth factor removal.

Example 10. Constructs for the treatment of neurodenerative disease

Constructs for treatment of neurodegenerative disease and containing the effectors SpiC, SptP or SopE2 were prepared as outlined in Example 9.

Example 11. Constructs for regulating cellular secretion and expression of cell surface receptors

For neuronal cells, constructs containing the effectors SpiC, SopE, RalF, SseE,F,G and J, PipA and B, SifA and B were prepared as outlined in Example 9.

For non-neuronal cells, the targeting domain may be replaced by a ligand with specificity for the target cell type. Such ligands may be selected from a list including: antibodies, carbohydrates, vitamins, hormones, cytokines, lectins, interleukins, peptides, growth factors, cell attachment proteins, toxin fragments, viral coat proteins.

Example 12 Constructs for the treatment of intracellular pathogens

Constructs containing the effectors SopE/SopE2, RalF, SpiC, SseE,F,G or J, PipA or B, SifA or B, or other effectors, for example those described in table 1, are useful therapeutic agents for treatment of disease.

Constructs were prepared essentially as described in example 9 but with a suitable binding domain selected from a list including; antibodies, carbohydrates, vitamins, hormones, cytokines, lectins, interleukins, peptides, growth factors, cell attachment proteins, toxin fragments, viral coat proteins etc. For targeting to macrophages this might include protective antigen from *Bacillus anthracis* or a carbohydrate moiety such as a mannose modification allowing specific uptake.

A recombinant construct of the invention includes an effector protein and a binding domain suitable for targeting the effector to a desired cell type.

When delivered to cells such constructs result in cellular events that cause the death of the intracellular pathogen, prevent its release from the infected cell type or otherwise reduce its ability to infect and cause disease.

Further embodiments of the invention are represented by all combinations of the recited effectors with the recited linker-translocation domain-binding domain conjugates.

The present application includes a sequence listing in which the following sequences referred to by their SEQ ID No.s represent the following embodiments of the invention:-

SEQ ID. NO.	DESCRIPTION
1	Diphtheria toxin translocation domain
2	Diphtheria toxin translocation domain, TeNT-Hc
3	Thrombin linker, Diphtheria toxin translocation domain, TeNT-Hc
4	Factor Xa linker, Diphtheria toxin translocation domain, TeNT-Hc
5	Diphtheria toxin translocation domain, BoNT/F-Hc
6	Thrombin linker, Diphtheria toxin translocation domain, BoNT/F-Hc
7	Factor Xa linker, Diphtheria toxin translocation domain, BoNT/F-Hc
8	AAC46234 invasion gene D protein [Salmonella typhimurium] SigD

9	AAF21057 invasion protein D [Salmonella typhimurium] SopB
10	CAC05808 IpgD, secreted by the Mxi-Spa machinery, modulates entry of bacteria into epithelial cells [Shigella flexneri]
11	AAC 69766 targeted effector protein [Yersinia pestis] YopJ
12	AAC02071 SopE [Salmonella typhimurium]
13	AAC44349 protein tyrosine phosphatase SptP [Salmonella typhimurium]
14	NP_047628 targeted effector [Yersinia pestis] YopE
15	AAK39624 exoenzyme S [Pseudomonas aeruginosa]
16	AAG03434 exoenzyme T [Pseudomonas aeruginosa]
17	NP_047619 Yop targeted effector [Yersinia pestis] YopT
18	NP_052380 protein kinase YopO [Yersinia enterocolitica]
19	AAF82095 outer protein AvrA [Salmonella enterica subsp. enterica serovar Dublin]
20	AAC44300 SpiC [Salmonella typhimurium]
21	SigD with the first 29 codons removed, thrombin linker, diphtheria translocation domain, TeNT-Hc
22	SigD with the first 29 codons removed, factor Xa linker, diphtheria translocation domain, TeNT-H_c
23	SigD with the first 29 codons removed, thrombin linker,

diphtheria toxin translocation domain, with BoNT/F-H_c

24 SigD, factor Xa linker, diphtheria toxin translocation domain, with BoNT/F-H_c

25 YopT, factor Xa linker, diphtheria translocation domain, TeNT-H_c

26 YopT, factor Xa linker, diphtheria toxin translocation domain, with BoNT/F-H_c

27 SpiC, thrombin linker, diphtheria translocation domain, TeNT-Hc

28 SpiC, factor Xa linker, diphtheria translocation domain, TeNT-H_c

29 SpiC fused to a domain consisting the N-terminal 254 residues from *Bacillus anthracis* lethal factor capable of interacting with protective antigen

30 *Bacillus anthracis* protective antigen

31 *Clostridium botulinum* C2 toxin component 1

32 *Clostridium botulinum* C2 toxin component 2

Table 1: Examples of type III and type IV effectors and their activity.

Effector	Biochemical function	Possible applications
YopT *Yersinia spp*	Inactivates Rho GTPases by direct	Stimulate nerve regrowth following damage
ExoS (N-terminal domain) *Pseudomonas aeuruginosa* YopE *Yersinia spp*	GTPase activating protein for Rho GTPases	Stimulate nerve regrowth
ExoS (C-terminal domain) *P.aeuruginosa*	ADP-ribosyltranferase modifies Ras and Rap GTPases	Block Ras/Rap signalling pathways
SptP (N-terminal domain) *Salmonella spp*	*GAP activity for Rac 1/ Cdc 42*	
SopE/E2 *S.typhimurium*	Guanine nucleotide exchange factor for Cdc42/Rac	Regulates nitric oxide release
YpkO/YopO *Yersinia spp*	Serine/threonine kinase modifies RhoA/Rac	
YopP/YopJ *Yersinia spp* AvrXv/AvrBsT *Xanthomonas campestris*	Blocks activation of various MAP kinase pathways	Induction of apoptosis in tumour cells Block release of inflammatory mediators from damaged cells
SopB/SigA/SigD *Salmonella spp* IpgD *Shigella flexneri*	Activate AKT kinase	Block apoptosis in damaged/ageing neurons
SpiC *S.enterica*	Block endosome fusion	Prevent neurotransmitter release from pain fibres
IpaB SipB	Induces apoptosis by direct activation of caspase 1	Induction of apoptosis in glioma/neuroblastoma cells
Orf19 *E.coli* IpgB *Shigella flexneri*	Affects mitochondrial function	Modulation of induction of cell death and other mitochondrial functions
Unidentified effector *Chlamydia spp*	Blocks apoptosis	Prevent apoptosis in damaged/ageing neurones
RalF *Listeria monocytogenes*	Guanine nucleotide exchange factor for ARF	Promote or prevent membrane compartment fusion
SpiC, SopE, SseE,F,G or J, PipA or B, SifA or B, *Salmonella spp.* RalF, *Listeria monocytogenes*	Various	Treating intracellular pathogens or disorders of intracellular trafficking
CagA *Helicobacter pylori*	Cytoskeletal modification	Alter uptake or release of membrane vesicle contents
YopM *Yersinia spp*, PopC *Ralstonia solaniacearum*	Leucine rich repeat protein. Possible transcription factors	Upregulation of genes involved in cell cycle and cell growth (YopM) or other genes.

CLAIMS

1. A conjugate of an injected bacterial effector protein and a carrier that targets the effector protein to a target cell.

2. A conjugate according to claim 1, comprising the effector protein linked by a linker to the carrier.

3. A conjugate according to claim 2, wherein the linker is cleavable, in that it can be cleaved after entry into the target cell so as to release the effector from the carrier.

4. A conjugate according to any previous claim, wherein the carrier targets the effector to a cell selected from an epithelial cell, a neuronal cell, a secretory cell, an immunological cell, an endocrine cell, an inflammatory cell, an exocrine cell, a bone cell and a cell of the cardiovascular system.

5. A conjugate according to any previous claim, wherein the carrier is a cell targeting component that comprises a first domain targeting the effector to a target cell and a second domain that translocates the effector into the cytosol of the cell.

6. A conjugate according to any previous claim, which is a single polypeptide.

7. A conjugate according to any previous claim, wherein the injected bacterial effector protein has an activity selected from activating GTPase, inactivating GTPase, enhancing replacement of bound GDP by GTP, causing covalent modification of GTPase, protein kinase activity, protein phosphatase, inositol phosphatase activity, inhibition of mitogen activated protein kinase kinase, regulation of gene expression, transcription factor and modulation of cellular trafficking.

8. A pharmaceutical composition comprising a conjugate according to any previous claim.

9. A pharmaceutical composition comprising a conjugate according to claims 1-7 for a treatment selected from promoting survival of cells, preventing damage to cells, reversing damage to cells, promoting growth of cells, inhibiting apoptosis, inhibiting release of an inflammatory mediator from cells, promoting division of cells and treating intracellular infection.

10. A pharmaceutical composition according to claims 8 or 9, for treating intracellular infection.

11. A pharmaceutical composition according to any of claims 8-10, for a treatment selected from inhibiting survival of cells, inhibiting growth of cells, inhibiting division of cells, promoting apoptosis, killing cells, promoting release of an inflammatory mediator from cells, regulating nitric oxide release from cells, inhibiting secretion from cells, interfering with intracellular trafficking and modulating expression of cell-surface markers.

12. A pharmaceutical composition according to claim 11, for interfering with intracellular trafficking.

13. A pharmaceutical composition according to claim 11, for modulating expression of cell-surface markers.

14. A pharmaceutical composition according to claim 11, for inhibiting secretion from cells

15. A pharmaceutical composition according to any of claims 8-14, for treatment of neuronal cells.

16. A pharmaceutical composition according to Claim 15, for promoting survival of neuronal cells.

17. A DNA construct encoding a conjugate according to any of claims 1-7.

18. A pharmaceutical composition, comprising the DNA construct of claim

17.

19. A pharmaceutical composition, comprising a vector containing the DNA construct of claim 17.

20. A pharmaceutical composition for delivery of an injected bacterial effector protein to a cell, comprising:-
the effector protein; linked by a cleavable linker to
a cell targeting component, comprising a first domain that binds to a cell and a second domain that translocates the effector protein of the composition into the cell.

21. A composition according to Claim 20 wherein the first domain is selected from (a) neuronal cell binding domains of clostridial toxins; and (b) fragments, variants and derivatives of the domains in (a) that substantially retain the neuronal cell binding activity of the domains of (a).

22. A composition according to Claim 20 or 21 wherein the second domain is selected from (a) domains of clostridial neurotoxins that translocate polypeptide sequences into cells, and (b) fragments, variants and derivatives of the domains of (a) that substantially retain the translocating activity of the domains of (a).

23. A composition according to Claim 20 or 21 wherein the second domain is selected from:-
(a) a translocation domain that is not a H_N domain of a clostridial toxin and is not a fragment or derivative of a H_N domain of a clostridial toxin;
(b) a non-aggregating translocation domain as measured by size in physiological buffers;
(c) a H_N domain of a diphtheria toxin,
(d) a fragment or derivative of (c) that substantially retains the translocating activity of the H_N domain of a diphtheria toxin,
(e) a fusogenic peptide,
(f) a membrane disrupting peptide, and
(g) translocating fragments and derivatives of (e) and (f).

24. A composition according to any of Claims 20 to 23 wherein the linker is cleaved in the neuronal cell so as to release the effector protein from the targeting component.

25. A composition according to Claim 24, wherein the linker is a disulphide bridge or a site for a protease found in the target cell.

26. A method of preparation of a conjugate according to any of claims 1 to 7 by combining the effector protein with the carrier.

27. A method according to claim 26, comprising chemically linking the effector protein with the carrier.

28. A method according to claim 26, comprising expressing a DNA that encodes a polypeptide having a first region that corresponds to the effector protein and a second region that codes for the carrier.

29. A method according to any of claims 26 to 28, wherein the polypeptide includes a third region, between the first and second regions, which is cleaved by a proteolytic enzyme present in the target cell.

30. A method according to any of claims 26 to 29, comprising linking the polypeptide between the first and second region and linking the first and second regions via a disulphide bridge.

31. Use of a conjugate of an injected bacterial effector protein and a carrier that targets the effector protein to a target cell in manufacture of a medicament.

32. Use of a DNA construct encoding a conjugate of an injected bacterial effector protein and a carrier that targets the effector protein to a target cell in manufacture of a medicament.

33. Use according to claims 31 or 32 in manufacture of a medicament for treatment of a neuronal cell.

34. Use according to claims 31 or 32 in manufacture of a medicament for treating intracellular infection.

35. Use according to claims 31 or 32 in manufacture of a medicament for interfering with intracellular trafficking.

36. Use according to claims 31 or 32 in manufacture of a medicament for modulating expression of cell-surface markers.

37. Use according to claims 31 or 32 in manufacture of a medicament for inhibiting secretion from cells.

38. Use of an injected bacterial effector protein in manufacture of a medicament for treatment of a neuronal cell.

39. Use of an injected bacterial effector protein in manufacture of a medicament for treating intracellular infection.

40. Use of an injected bacterial effector protein in manufacture of a medicament for interfering with intracellular trafficking.

41. Use of an injected bacterial type effector protein in manufacture of a medicament for modulating expression of cell-surface markers.

42. Use of an injected bacterial effector protein in manufacture of a medicament for inhibiting secretion from cells.

43. Use of a DNA construct encoding an injected bacterial effector protein in manufacture of a medicament for treatment of a neuronal cell.

44. Use of a DNA construct encoding an injected bacterial effector protein in manufacture of a medicament for treating intracellular infection.

45. Use of a DNA construct encoding an injected bacterial effector protein in manufacture of a medicament for interfering with

intracellular trafficking.

46. Use of a DNA construct encoding an injected bacterial effector protein in manufacture of a medicament for modulating expression of cell-surface markers.

47. Use of a DNA construct encoding an injected bacterial effector protein in manufacture of a medicament for inhibiting secretion from cells.

48. A method of delivering an injected bacterial effector protein to a neuronal cell comprising administering a composition according to any of Claims 19 to 24.

49. A method according to Claim 46 comprising injecting the composition.

SEQUENCE LISTING

<110> Microbiological Research Authority

Clifford, Shone C

John, Sutton M

Nigel, Silman

<120> Pharmaceutical use of secreted bacterial effector proteins

<130> GWS/PG/23433

<160> 32

<170> PatentIn version 3.1

<210> 1

<211> 210

<212> PRT

<213> diphtheria toxin translocation domain

<400> 1

Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp Val Ile
1 5 10 15

Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly Pro
20 25 30

Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu Glu
35 40 45

Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu His
50 55 60

Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val Phe
65 70 75 80

Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val Ile
85 90 95

Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu Ser
100 105 110

Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala Val
115 120 125

His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu Ser Ser
130 135 140

Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp Ile
145 150 155 160

Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe Gln
165 170 175

Val Val His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly His Lys
180 185 190

Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr Val
195 200 205

Arg Ser
210

<210> 2

<211> 665

<212> PRT

<213> diphtheria toxin translocation domain TeNT-HC

<400> 2

Gly Ser Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp
1 5 10 15

Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His
20 25 30

Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser
35 40 45

Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu
50 55 60

Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro
65 70 75 80

Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln

```
                85                  90                  95

Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala
            100                 105                 110

Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly
        115                 120                 125

Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu
    130                 135                 140

Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val
145                 150                 155                 160

Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu
                165                 170                 175

Phe Gln Val Val His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly
            180                 185                 190

His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn
        195                 200                 205

Thr Val Arg Ser Lys Asn Leu Asp Cys Trp Val Asp Asn Glu Glu Asp
    210                 215                 220

Ile Asp Val Ile Leu Lys Lys Ser Thr Ile Leu Asn Leu Asp Ile Asn
225                 230                 235                 240

Asn Asp Ile Ile Ser Asp Ile Ser Gly Phe Asn Ser Ser Val Ile Thr
                245                 250                 255

Tyr Pro Asp Ala Gln Leu Val Pro Gly Ile Asn Gly Lys Ala Ile His
            260                 265                 270

Leu Val Asn Asn Glu Ser Ser Glu Val Ile Val His Lys Ala Met Asp
        275                 280                 285

Ile Glu Tyr Asn Asp Met Phe Asn Asn Phe Thr Val Ser Phe Trp Leu
    290                 295                 300

Arg Val Pro Lys Val Ser Ala Ser His Leu Glu Gln Tyr Gly Thr Asn
305                 310                 315                 320

Glu Tyr Ser Ile Ile Ser Ser Met Lys Lys His Ser Leu Ser Ile Gly
                325                 330                 335

Ser Gly Trp Ser Val Ser Leu Lys Gly Asn Asn Leu Ile Trp Thr Leu
            340                 345                 350
```

Lys Asp Ser Ala Gly Glu Val Arg Gln Ile Thr Phe Arg Asp Leu Pro
355 360 365

Asp Lys Phe Asn Ala Tyr Leu Ala Asn Lys Trp Val Phe Ile Thr Ile
370 375 380

Thr Asn Asp Arg Leu Ser Ser Ala Asn Leu Tyr Ile Asn Gly Val Leu
385 390 395 400

Met Gly Ser Ala Glu Ile Thr Gly Leu Gly Ala Ile Arg Glu Asp Asn
405 410 415

Asn Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn Asn Asn Gln Tyr Val
420 425 430

Ser Ile Asp Lys Phe Arg Ile Phe Cys Lys Ala Leu Asn Pro Lys Glu
435 440 445

Ile Glu Lys Leu Tyr Thr Ser Tyr Leu Ser Ile Thr Phe Leu Arg Asp
450 455 460

Phe Trp Gly Asn Pro Leu Arg Tyr Asp Thr Glu Tyr Tyr Leu Ile Pro
465 470 475 480

Val Ala Ser Ser Ser Lys Asp Val Gln Leu Lys Asn Ile Thr Asp Tyr
485 490 495

Met Tyr Leu Thr Asn Ala Pro Ser Tyr Thr Asn Gly Lys Leu Asn Ile
500 505 510

Tyr Tyr Arg Arg Leu Tyr Asn Gly Leu Lys Phe Ile Ile Lys Arg Tyr
515 520 525

Thr Pro Asn Asn Glu Ile Asp Ser Phe Val Lys Ser Gly Asp Phe Ile
530 535 540

Lys Leu Tyr Val Ser Tyr Asn Asn Asn Glu His Ile Val Gly Tyr Pro
545 550 555 560

Lys Asp Gly Asn Ala Phe Asn Asn Leu Asp Arg Ile Leu Arg Val Gly
565 570 575

Tyr Asn Ala Pro Gly Ile Pro Leu Tyr Lys Lys Met Glu Ala Val Lys
580 585 590

Leu Arg Asp Leu Lys Thr Tyr Ser Val Gln Leu Lys Leu Tyr Asp Asp
595 600 605

Lys Asn Ala Ser Leu Gly Leu Val Gly Thr His Asn Gly Gln Ile Gly
610 615 620

Asn Asp Pro Asn Arg Asp Ile Leu Ile Ala Ser Asn Trp Tyr Phe Asn
625 630 635 640

His Leu Lys Asp Lys Ile Leu Gly Cys Asp Trp Tyr Phe Val Pro Thr
645 650 655

Asp Glu Gly Trp Thr Asn Asp Leu Gln
660 665

<210> 3

<211> 677

<212> PRT

<213> thrombin linker, diphtheria toxin translocation domain, TeNT-HC

<400> 3

Arg Ser Cys Gly Leu Val Pro Arg Gly Ser Gly Pro Gly Ser Ser Val
1 5 10 15

Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp Val Ile Arg Asp
20 25 30

Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly Pro Ile Lys
35 40 45

Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu Glu Lys Ala
50 55 60

Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu His Pro Glu
65 70 75 80

Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val Phe Ala Gly
85 90 95

Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val Ile Asp Ser
100 105 110

Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu Ser Ile Leu
115 120 125

Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala Val His His
130 135 140

Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu Ser Ser Leu Met
145 150 155 160

Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp Ile Gly Phe

```
                165                 170                 175

Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe Gln Val Val
            180                 185                 190

His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly His Lys Thr Gln
        195                 200                 205

Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr Val Arg Ser
    210                 215                 220

Lys Asn Leu Asp Cys Trp Val Asp Asn Glu Glu Asp Ile Asp Val Ile
225                 230                 235                 240

Leu Lys Lys Ser Thr Ile Leu Asn Leu Asp Ile Asn Asn Asp Ile Ile
                245                 250                 255

Ser Asp Ile Ser Gly Phe Asn Ser Ser Val Ile Thr Tyr Pro Asp Ala
            260                 265                 270

Gln Leu Val Pro Gly Ile Asn Gly Lys Ala Ile His Leu Val Asn Asn
        275                 280                 285

Glu Ser Ser Glu Val Ile Val His Lys Ala Met Asp Ile Glu Tyr Asn
    290                 295                 300

Asp Met Phe Asn Asn Phe Thr Val Ser Phe Trp Leu Arg Val Pro Lys
305                 310                 315                 320

Val Ser Ala Ser His Leu Glu Gln Tyr Gly Thr Asn Glu Tyr Ser Ile
                325                 330                 335

Ile Ser Ser Met Lys Lys His Ser Leu Ser Ile Gly Ser Gly Trp Ser
            340                 345                 350

Val Ser Leu Lys Gly Asn Asn Leu Ile Trp Thr Leu Lys Asp Ser Ala
        355                 360                 365

Gly Glu Val Arg Gln Ile Thr Phe Arg Asp Leu Pro Asp Lys Phe Asn
    370                 375                 380

Ala Tyr Leu Ala Asn Lys Trp Val Phe Ile Thr Ile Thr Asn Asp Arg
385                 390                 395                 400

Leu Ser Ser Ala Asn Leu Tyr Ile Asn Gly Val Leu Met Gly Ser Ala
                405                 410                 415

Glu Ile Thr Gly Leu Gly Ala Ile Arg Glu Asp Asn Asn Ile Thr Leu
            420                 425                 430
```

Lys Leu Asp Arg Cys Asn Asn Asn Asn Gln Tyr Val Ser Ile Asp Lys
435 440 445

Phe Arg Ile Phe Cys Lys Ala Leu Asn Pro Lys Glu Ile Glu Lys Leu
450 455 460

Tyr Thr Ser Tyr Leu Ser Ile Thr Phe Leu Arg Asp Phe Trp Gly Asn
465 470 475 480

Pro Leu Arg Tyr Asp Thr Glu Tyr Tyr Leu Ile Pro Val Ala Ser Ser
485 490 495

Ser Lys Asp Val Gln Leu Lys Asn Ile Thr Asp Tyr Met Tyr Leu Thr
500 505 510

Asn Ala Pro Ser Tyr Thr Asn Gly Lys Leu Asn Ile Tyr Tyr Arg Arg
515 520 525

Leu Tyr Asn Gly Leu Lys Phe Ile Ile Lys Arg Tyr Thr Pro Asn Asn
530 535 540

Glu Ile Asp Ser Phe Val Lys Ser Gly Asp Phe Ile Lys Leu Tyr Val
545 550 555 560

Ser Tyr Asn Asn Asn Glu His Ile Val Gly Tyr Pro Lys Asp Gly Asn
565 570 575

Ala Phe Asn Asn Leu Asp Arg Ile Leu Arg Val Gly Tyr Asn Ala Pro
580 585 590

Gly Ile Pro Leu Tyr Lys Lys Met Glu Ala Val Lys Leu Arg Asp Leu
595 600 605

Lys Thr Tyr Ser Val Gln Leu Lys Leu Tyr Asp Asp Lys Asn Ala Ser
610 615 620

Leu Gly Leu Val Gly Thr His Asn Gly Gln Ile Gly Asn Asp Pro Asn
625 630 635 640

Arg Asp Ile Leu Ile Ala Ser Asn Trp Tyr Phe Asn His Leu Lys Asp
645 650 655

Lys Ile Leu Gly Cys Asp Trp Tyr Phe Val Pro Thr Asp Glu Gly Trp
660 665 670

Thr Asn Asp Leu Gln
675

<210> 4

<211> 677

<212> PRT

<213> factor Xa linker, Diphtheria toxin translocation domain, TeNT-HC

<400> 4

Arg Ser Cys Gly Ile Glu Gly Arg Ala Pro Gly Pro Gly Ser Ser Val
1 5 10 15

Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp Val Ile Arg Asp
20 25 30

Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly Pro Ile Lys
35 40 45

Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu Glu Lys Ala
50 55 60

Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu His Pro Glu
65 70 75 80

Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val Phe Ala Gly
85 90 95

Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val Ile Asp Ser
100 105 110

Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu Ser Ile Leu
115 120 125

Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala Val His His
130 135 140

Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu Ser Ser Leu Met
145 150 155 160

Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp Ile Gly Phe
165 170 175

Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe Gln Val Val
180 185 190

His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly His Lys Thr Gln
195 200 205

Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr Val Arg Ser
210 215 220

```
Lys Asn Leu Asp Cys Trp Val Asp Asn Glu Glu Asp Ile Asp Val Ile
225                 230                 235                 240

Leu Lys Lys Ser Thr Ile Leu Asn Leu Asp Ile Asn Asn Asp Ile Ile
                245                 250                 255

Ser Asp Ile Ser Gly Phe Asn Ser Ser Val Ile Thr Tyr Pro Asp Ala
            260                 265                 270

Gln Leu Val Pro Gly Ile Asn Gly Lys Ala Ile His Leu Val Asn Asn
        275                 280                 285

Glu Ser Ser Glu Val Ile Val His Lys Ala Met Asp Ile Glu Tyr Asn
    290                 295                 300

Asp Met Phe Asn Asn Phe Thr Val Ser Phe Trp Leu Arg Val Pro Lys
305                 310                 315                 320

Val Ser Ala Ser His Leu Glu Gln Tyr Gly Thr Asn Glu Tyr Ser Ile
                325                 330                 335

Ile Ser Ser Met Lys Lys His Ser Leu Ser Ile Gly Ser Gly Trp Ser
            340                 345                 350

Val Ser Leu Lys Gly Asn Asn Leu Ile Trp Thr Leu Lys Asp Ser Ala
        355                 360                 365

Gly Glu Val Arg Gln Ile Thr Phe Arg Asp Leu Pro Asp Lys Phe Asn
    370                 375                 380

Ala Tyr Leu Ala Asn Lys Trp Val Phe Ile Thr Ile Thr Asn Asp Arg
385                 390                 395                 400

Leu Ser Ser Ala Asn Leu Tyr Ile Asn Gly Val Leu Met Gly Ser Ala
                405                 410                 415

Glu Ile Thr Gly Leu Gly Ala Ile Arg Glu Asp Asn Asn Ile Thr Leu
            420                 425                 430

Lys Leu Asp Arg Cys Asn Asn Asn Asn Gln Tyr Val Ser Ile Asp Lys
        435                 440                 445

Phe Arg Ile Phe Cys Lys Ala Leu Asn Pro Lys Glu Ile Glu Lys Leu
    450                 455                 460

Tyr Thr Ser Tyr Leu Ser Ile Thr Phe Leu Arg Asp Phe Trp Gly Asn
465                 470                 475                 480

Pro Leu Arg Tyr Asp Thr Glu Tyr Tyr Leu Ile Pro Val Ala Ser Ser
                485                 490                 495
```

Ser Lys Asp Val Gln Leu Lys Asn Ile Thr Asp Tyr Met Tyr Leu Thr
500 505 510

Asn Ala Pro Ser Tyr Thr Asn Gly Lys Leu Asn Ile Tyr Tyr Arg Arg
515 520 525

Leu Tyr Asn Gly Leu Lys Phe Ile Ile Lys Arg Tyr Thr Pro Asn Asn
530 535 540

Glu Ile Asp Ser Phe Val Lys Ser Gly Asp Phe Ile Lys Leu Tyr Val
545 550 555 560

Ser Tyr Asn Asn Asn Glu His Ile Val Gly Tyr Pro Lys Asp Gly Asn
565 570 575

Ala Phe Asn Asn Leu Asp Arg Ile Leu Arg Val Gly Tyr Asn Ala Pro
580 585 590

Gly Ile Pro Leu Tyr Lys Lys Met Glu Ala Val Lys Leu Arg Asp Leu
595 600 605

Lys Thr Tyr Ser Val Gln Leu Lys Leu Tyr Asp Asp Lys Asn Ala Ser
610 615 620

Leu Gly Leu Val Gly Thr His Asn Gly Gln Ile Gly Asn Asp Pro Asn
625 630 635 640

Arg Asp Ile Leu Ile Ala Ser Asn Trp Tyr Phe Asn His Leu Lys Asp
645 650 655

Lys Ile Leu Gly Cys Asp Trp Tyr Phe Val Pro Thr Asp Glu Gly Trp
660 665 670

Thr Asn Asp Leu Gln
675

<210> 5

<211> 645

<212> PRT

<213> diphtheria toxin translocation domain with BoNT/F-HC

<400> 5

Gly Ser Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp
1 5 10 15

Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His

```
            20                  25                  30

Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser
        35                  40                  45

Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu
    50                  55                  60

Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro
65                  70                  75                  80

Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln
                85                  90                  95

Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala
            100                 105                 110

Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly
        115                 120                 125

Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu
    130                 135                 140

Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val
145                 150                 155                 160

Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu
                165                 170                 175

Phe Gln Val Val His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly
            180                 185                 190

His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn
        195                 200                 205

Thr Val Arg Ser Thr Met Ser Tyr Thr Asn Asp Lys Ile Leu Ile Leu
    210                 215                 220

Tyr Phe Asn Lys Leu Tyr Lys Lys Ile Lys Asp Asn Ser Ile Leu Asp
225                 230                 235                 240

Met Arg Tyr Glu Asn Asn Lys Phe Ile Asp Ile Ser Gly Tyr Gly Ser
                245                 250                 255

Asn Ile Ser Ile Asn Gly Asp Val Tyr Ile Tyr Ser Thr Asn Arg Asn
            260                 265                 270

Gln Phe Gly Ile Tyr Ser Ser Lys Pro Ser Glu Val Asn Ile Ala Gln
        275                 280                 285
```

```
Asn Asn Asp Ile Ile Tyr Asn Gly Arg Tyr Gln Asn Phe Ser Ile Ser
    290                 295                 300

Phe Trp Val Arg Ile Pro Lys Tyr Phe Asn Lys Val Asn Leu Asn Asn
305                 310                 315                 320

Glu Tyr Thr Ile Ile Asp Cys Ile Arg Asn Asn Asn Ser Gly Trp Lys
                325                 330                 335

Ile Ser Leu Asn Tyr Asn Lys Ile Ile Trp Thr Leu Gln Asp Thr Ala
            340                 345                 350

Gly Asn Asn Gln Lys Leu Val Phe Asn Tyr Thr Gln Met Ile Ser Ile
        355                 360                 365

Ser Asp Tyr Ile Asn Lys Trp Ile Phe Val Thr Ile Thr Asn Asn Arg
    370                 375                 380

Leu Gly Asn Ser Arg Ile Tyr Ile Asn Gly Asn Leu Ile Asp Glu Lys
385                 390                 395                 400

Ser Ile Ser Asn Leu Gly Asp Ile His Val Ser Asp Asn Ile Leu Phe
                405                 410                 415

Lys Ile Val Gly Cys Asn Asp Thr Arg Tyr Val Gly Ile Arg Tyr Phe
            420                 425                 430

Lys Val Phe Asp Thr Glu Leu Gly Lys Thr Glu Ile Glu Thr Leu Tyr
        435                 440                 445

Ser Asp Glu Pro Asp Pro Ser Ile Leu Lys Asp Phe Trp Gly Asn Tyr
    450                 455                 460

Leu Leu Tyr Asn Lys Arg Tyr Tyr Leu Leu Asn Leu Leu Arg Thr Asp
465                 470                 475                 480

Lys Ser Ile Thr Gln Asn Ser Asn Phe Leu Asn Ile Asn Gln Gln Arg
                485                 490                 495

Gly Val Tyr Gln Lys Pro Asn Ile Phe Ser Asn Thr Arg Leu Tyr Thr
            500                 505                 510

Gly Val Glu Val Ile Ile Arg Lys Asn Gly Ser Thr Asp Ile Ser Asn
        515                 520                 525

Thr Asp Asn Phe Val Arg Lys Asn Asp Leu Ala Tyr Ile Asn Val Val
    530                 535                 540

Asp Arg Asp Val Glu Tyr Arg Leu Tyr Ala Asp Ile Ser Ile Ala Lys
545                 550                 555                 560
```

Pro Glu Lys Ile Ile Lys Leu Ile Arg Thr Ser Asn Ser Asn Asn Ser
565 570 575

Leu Gly Gln Ile Ile Val Met Asp Ser Ile Gly Asn Asn Cys Thr Met
580 585 590

Asn Phe Gln Asn Asn Asn Gly Gly Asn Ile Gly Leu Leu Gly Phe His
595 600 605

Ser Asn Asn Leu Val Ala Ser Ser Trp Tyr Tyr Asn Asn Ile Arg Lys
610 615 620

Asn Thr Ser Ser Asn Gly Cys Phe Trp Ser Phe Ile Ser Lys Glu His
625 630 635 640

Gly Trp Gln Glu Asn
645

<210> 6

<211> 657

<212> PRT

<213> thrombin linker, diphtheria toxin translocation domain, BoNT/F-HC

<400> 6

Arg Ser Cys Gly Leu Val Pro Arg Gly Ser Gly Pro Gly Ser Ser Val
1 5 10 15

Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp Val Ile Arg Asp
20 25 30

Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly Pro Ile Lys
35 40 45

Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu Glu Lys Ala
50 55 60

Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu His Pro Glu
65 70 75 80

Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val Phe Ala Gly
85 90 95

Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val Ile Asp Ser
100 105 110

Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu Ser Ile Leu

115 120 125

Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala Val His His
130 135 140

Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu Ser Ser Leu Met
145 150 155 160

Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp Ile Gly Phe
165 170 175

Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe Gln Val Val
180 185 190

His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly His Lys Thr Gln
195 200 205

Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr Val Arg Ser
210 215 220

Thr Met Ser Tyr Thr Asn Asp Lys Ile Leu Ile Leu Tyr Phe Asn Lys
225 230 235 240

Leu Tyr Lys Lys Ile Lys Asp Asn Ser Ile Leu Asp Met Arg Tyr Glu
245 250 255

Asn Asn Lys Phe Ile Asp Ile Ser Gly Tyr Gly Ser Asn Ile Ser Ile
260 265 270

Asn Gly Asp Val Tyr Ile Tyr Ser Thr Asn Arg Asn Gln Phe Gly Ile
275 280 285

Tyr Ser Ser Lys Pro Ser Glu Val Asn Ile Ala Gln Asn Asn Asp Ile
290 295 300

Ile Tyr Asn Gly Arg Tyr Gln Asn Phe Ser Ile Ser Phe Trp Val Arg
305 310 315 320

Ile Pro Lys Tyr Phe Asn Lys Val Asn Leu Asn Asn Glu Tyr Thr Ile
325 330 335

Ile Asp Cys Ile Arg Asn Asn Asn Ser Gly Trp Lys Ile Ser Leu Asn
340 345 350

Tyr Asn Lys Ile Ile Trp Thr Leu Gln Asp Thr Ala Gly Asn Asn Gln
355 360 365

Lys Leu Val Phe Asn Tyr Thr Gln Met Ile Ser Ile Ser Asp Tyr Ile
370 375 380

Asn Lys Trp Ile Phe Val Thr Ile Thr Asn Asn Arg Leu Gly Asn Ser
385 390 395 400

Arg Ile Tyr Ile Asn Gly Asn Leu Ile Asp Glu Lys Ser Ile Ser Asn
405 410 415

Leu Gly Asp Ile His Val Ser Asp Asn Ile Leu Phe Lys Ile Val Gly
420 425 430

Cys Asn Asp Thr Arg Tyr Val Gly Ile Arg Tyr Phe Lys Val Phe Asp
435 440 445

Thr Glu Leu Gly Lys Thr Glu Ile Glu Thr Leu Tyr Ser Asp Glu Pro
450 455 460

Asp Pro Ser Ile Leu Lys Asp Phe Trp Gly Asn Tyr Leu Leu Tyr Asn
465 470 475 480

Lys Arg Tyr Tyr Leu Leu Asn Leu Leu Arg Thr Asp Lys Ser Ile Thr
485 490 495

Gln Asn Ser Asn Phe Leu Asn Ile Asn Gln Gln Arg Gly Val Tyr Gln
500 505 510

Lys Pro Asn Ile Phe Ser Asn Thr Arg Leu Tyr Thr Gly Val Glu Val
515 520 525

Ile Ile Arg Lys Asn Gly Ser Thr Asp Ile Ser Asn Thr Asp Asn Phe
530 535 540

Val Arg Lys Asn Asp Leu Ala Tyr Ile Asn Val Val Asp Arg Asp Val
545 550 555 560

Glu Tyr Arg Leu Tyr Ala Asp Ile Ser Ile Ala Lys Pro Glu Lys Ile
565 570 575

Ile Lys Leu Ile Arg Thr Ser Asn Ser Asn Asn Ser Leu Gly Gln Ile
580 585 590

Ile Val Met Asp Ser Ile Gly Asn Asn Cys Thr Met Asn Phe Gln Asn
595 600 605

Asn Asn Gly Gly Asn Ile Gly Leu Leu Gly Phe His Ser Asn Asn Leu
610 615 620

Val Ala Ser Ser Trp Tyr Tyr Asn Asn Ile Arg Lys Asn Thr Ser Ser
625 630 635 640

Asn Gly Cys Phe Trp Ser Phe Ile Ser Lys Glu His Gly Trp Gln Glu
645 650 655

Asn

<210> 7

<211> 657

<212> PRT

<213> factor Xa linker, diphtheria toxin translocation domain, BoNT/F-Hc

<400> 7

Arg Ser Cys Gly Ile Glu Gly Arg Ala Pro Gly Pro Gly Ser Ser Val
1 5 10 15

Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp Val Ile Arg Asp
20 25 30

Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly Pro Ile Lys
35 40 45

Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu Glu Lys Ala
50 55 60

Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu His Pro Glu
65 70 75 80

Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val Phe Ala Gly
85 90 95

Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val Ile Asp Ser
100 105 110

Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu Ser Ile Leu
115 120 125

Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala Val His His
130 135 140

Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu Ser Ser Leu Met
145 150 155 160

Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp Ile Gly Phe
165 170 175

Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe Gln Val Val
180 185 190

His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly His Lys Thr Gln
195 200 205

Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr Val Arg Ser
210 215 220

Thr Met Ser Tyr Thr Asn Asp Lys Ile Leu Ile Leu Tyr Phe Asn Lys
225 230 235 240

Leu Tyr Lys Lys Ile Lys Asp Asn Ser Ile Leu Asp Met Arg Tyr Glu
245 250 255

Asn Asn Lys Phe Ile Asp Ile Ser Gly Tyr Gly Ser Asn Ile Ser Ile
260 265 270

Asn Gly Asp Val Tyr Ile Tyr Ser Thr Asn Arg Asn Gln Phe Gly Ile
275 280 285

Tyr Ser Ser Lys Pro Ser Glu Val Asn Ile Ala Gln Asn Asn Asp Ile
290 295 300

Ile Tyr Asn Gly Arg Tyr Gln Asn Phe Ser Ile Ser Phe Trp Val Arg
305 310 315 320

Ile Pro Lys Tyr Phe Asn Lys Val Asn Leu Asn Asn Glu Tyr Thr Ile
325 330 335

Ile Asp Cys Ile Arg Asn Asn Asn Ser Gly Trp Lys Ile Ser Leu Asn
340 345 350

Tyr Asn Lys Ile Ile Trp Thr Leu Gln Asp Thr Ala Gly Asn Asn Gln
355 360 365

Lys Leu Val Phe Asn Tyr Thr Gln Met Ile Ser Ile Ser Asp Tyr Ile
370 375 380

Asn Lys Trp Ile Phe Val Thr Ile Thr Asn Asn Arg Leu Gly Asn Ser
385 390 395 400

Arg Ile Tyr Ile Asn Gly Asn Leu Ile Asp Glu Lys Ser Ile Ser Asn
405 410 415

Leu Gly Asp Ile His Val Ser Asp Asn Ile Leu Phe Lys Ile Val Gly
420 425 430

Cys Asn Asp Thr Arg Tyr Val Gly Ile Arg Tyr Phe Lys Val Phe Asp
435 440 445

Thr Glu Leu Gly Lys Thr Glu Ile Glu Thr Leu Tyr Ser Asp Glu Pro
450 455 460

Asp Pro Ser Ile Leu Lys Asp Phe Trp Gly Asn Tyr Leu Leu Tyr Asn
465 470 475 480

Lys Arg Tyr Tyr Leu Leu Asn Leu Leu Arg Thr Asp Lys Ser Ile Thr
485 490 495

Gln Asn Ser Asn Phe Leu Asn Ile Asn Gln Gln Arg Gly Val Tyr Gln
500 505 510

Lys Pro Asn Ile Phe Ser Asn Thr Arg Leu Tyr Thr Gly Val Glu Val
515 520 525

Ile Ile Arg Lys Asn Gly Ser Thr Asp Ile Ser Asn Thr Asp Asn Phe
530 535 540

Val Arg Lys Asn Asp Leu Ala Tyr Ile Asn Val Val Asp Arg Asp Val
545 550 555 560

Glu Tyr Arg Leu Tyr Ala Asp Ile Ser Ile Ala Lys Pro Glu Lys Ile
565 570 575

Ile Lys Leu Ile Arg Thr Ser Asn Ser Asn Asn Ser Leu Gly Gln Ile
580 585 590

Ile Val Met Asp Ser Ile Gly Asn Asn Cys Thr Met Asn Phe Gln Asn
595 600 605

Asn Asn Gly Gly Asn Ile Gly Leu Leu Gly Phe His Ser Asn Asn Leu
610 615 620

Val Ala Ser Ser Trp Tyr Tyr Asn Asn Ile Arg Lys Asn Thr Ser Ser
625 630 635 640

Asn Gly Cys Phe Trp Ser Phe Ile Ser Lys Glu His Gly Trp Gln Glu
645 650 655

Asn

<210> 8

<211> 563

<212> PRT

<213> AAC46234 invasion gene D protein [Salmonella typhimurium] SigD

<400> 8

Met Gln Ile Gln Ser Phe Tyr His Ser Ala Ser Leu Lys Thr Gln Glu

```
1               5                   10                  15

Ala Phe Lys Ser Leu Gln Lys Thr Leu Tyr Asn Gly Met Gln Ile Leu
            20                  25                  30

Ser Gly Gln Gly Lys Ala Pro Ala Lys Ala Pro Asp Ala Arg Pro Glu
        35                  40                  45

Ile Ile Val Leu Arg Glu Pro Gly Ala Thr Trp Gly Asn Tyr Leu Gln
    50                  55                  60

His Gln Lys Ala Ser Asn His Ser Leu His Asn Leu Tyr Asn Leu Gln
65                  70                  75                  80

Arg Asp Leu Leu Thr Val Ala Ala Thr Val Leu Gly Lys Gln Asp Pro
                85                  90                  95

Val Leu Thr Ser Met Ala Asn Gln Met Glu Leu Ala Lys Val Lys Ala
            100                 105                 110

Asp Arg Pro Ala Thr Lys Gln Glu Glu Ala Ala Ala Lys Ala Leu Lys
        115                 120                 125

Lys Asn Leu Ile Glu Leu Ile Ala Ala Arg Thr Gln Gln Gln Asp Gly
    130                 135                 140

Leu Pro Ala Lys Glu Ala His Arg Phe Ala Ala Val Ala Phe Arg Asp
145                 150                 155                 160

Ala Gln Val Lys Gln Leu Asn Asn Gln Pro Trp Gln Thr Ile Lys Asn
                165                 170                 175

Thr Leu Thr His Asn Gly His His Tyr Thr Asn Thr Gln Leu Pro Ala
            180                 185                 190

Ala Glu Met Lys Ile Gly Ala Lys Asp Ile Phe Pro Ser Ala Tyr Glu
        195                 200                 205

Gly Lys Gly Val Cys Ser Trp Asp Thr Lys Asn Ile His His Ala Asn
    210                 215                 220

Asn Leu Trp Met Ser Thr Val Ser Val His Glu Asp Gly Lys Asp Lys
225                 230                 235                 240

Thr Leu Phe Phe Asp Gly Ile Arg His Gly Val Leu Ser Pro Tyr His
                245                 250                 255

Glu Lys Asp Pro Leu Leu Arg His Val Gly Ala Glu Asn Lys Ala Lys
            260                 265                 270
```

Glu Val Leu Thr Ala Ala Leu Phe Ser Lys Pro Glu Leu Leu Asn Lys
275 280 285

Ala Leu Ala Gly Glu Ala Val Ser Leu Lys Leu Val Ser Val Gly Leu
290 295 300

Leu Thr Ala Ser Asn Ile Phe Gly Lys Glu Gly Thr Met Val Glu Asp
305 310 315 320

Gln Met Arg Ala Trp Gln Ser Leu Thr Gln Pro Gly Lys Met Ile His
325 330 335

Leu Lys Ile Arg Asn Lys Asp Gly Asp Leu Gln Thr Val Lys Ile Lys
340 345 350

Pro Asp Val Val Ala Ala Phe Asn Val Gly Val Asn Glu Leu Ala Leu
355 360 365

Lys Leu Gly Phe Gly Leu Lys Ala Ser Asp Ser Tyr Asn Ala Glu Ala
370 375 380

Leu His Gln Leu Leu Gly Asn Asp Leu Arg Pro Glu Ala Arg Pro Gly
385 390 395 400

Gly Trp Val Gly Glu Trp Leu Ala Gln Tyr Pro Asp Asn Tyr Glu Val
405 410 415

Val Asn Thr Leu Ala Arg Gln Ile Lys Asp Ile Trp Lys Asn Asn Gln
420 425 430

His His Lys Asp Gly Gly Glu Pro Tyr Lys Leu Ala Gln Arg Leu Ala
435 440 445

Met Leu Ala His Glu Ile Asp Ala Val Pro Ala Trp Asn Cys Lys Ser
450 455 460

Gly Lys Asp Arg Thr Gly Met Met Asp Ser Glu Ile Lys Gly Glu Ile
465 470 475 480

Ile Ser Leu His Gln Thr His Met Leu Ser Ala Pro Gly Ser Leu Pro
485 490 495

Asp Ser Gly Gly Gln Lys Ile Phe Gln Lys Val Leu Leu Asn Ser Gly
500 505 510

Asn Leu Glu Ile Gln Lys Gln Asn Thr Gly Gly Ala Gly Asn Lys Val
515 520 525

Met Lys Asn Leu Ser Pro Glu Val Leu Asn Leu Ser Tyr Gln Lys Arg
530 535 540

Val Gly Asp Glu Asn Ile Trp Gln Ser Val Lys Gly Ile Ser Ser Leu
545 550 555 560

Ile Thr Ser

<210> 9

<211> 433

<212> PRT

<213> AAF21057 invasion protein D [Salmonella typhimurium] SopB

<400> 9

Val Leu Thr Ser Met Ala Asn Gln Met Glu Leu Ala Lys Val Lys Ala
1 5 10 15

Asp Arg Pro Ala Thr Lys Gln Glu Glu Ala Ala Ala Lys Ala Leu Lys
20 25 30

Lys Asn Leu Ile Glu Leu Ile Ala Ala Arg Thr Gln Gln Gln Asp Gly
35 40 45

Leu Pro Ala Lys Glu Ala His Arg Phe Ala Ala Val Ala Phe Arg Asp
50 55 60

Ala Gln Val Lys Gln Leu Asn Asn Gln Pro Trp Gln Thr Ile Lys Asn
65 70 75 80

Thr Leu Thr His Asn Gly His His Tyr Thr Asn Thr Gln Leu Pro Ala
85 90 95

Ala Glu Met Lys Ile Gly Ala Lys Asp Ile Phe Pro Ser Ala Tyr Glu
100 105 110

Gly Lys Gly Val Cys Ser Trp Asp Thr Lys Asn Ile His His Ala Asn
115 120 125

Asn Leu Trp Met Ser Thr Val Ser Val His Glu Asp Gly Lys Asp Lys
130 135 140

Thr Leu Phe Cys Gly Ile Arg His Gly Val Leu Ser Pro Tyr His Glu
145 150 155 160

Lys Asp Pro Leu Leu Arg His Val Gly Ala Glu Asn Lys Ala Lys Glu
165 170 175

Val Leu Thr Ala Ala Leu Phe Ser Lys Pro Glu Leu Leu Asn Lys Ala

180 185 190

Leu Ala Gly Glu Ala Val Ser Leu Lys Leu Val Ser Val Gly Leu Leu
195 200 205

Thr Ala Ser Asn Ile Phe Gly Lys Glu Gly Thr Met Val Glu Asp Gln
210 215 220

Met Arg Ala Trp Gln Ser Leu Thr Gln Pro Gly Lys Met Ile His Leu
225 230 235 240

Lys Ile Arg Asn Lys Asp Gly Asp Leu Gln Thr Val Lys Ile Lys Pro
245 250 255

Asp Val Ala Ala Phe Asn Val Gly Val Asn Glu Leu Ala Leu Lys Leu
260 265 270

Gly Phe Gly Leu Lys Ala Ser Asp Ser Tyr Asn Ala Glu Ala Leu His
275 280 285

Gln Leu Leu Gly Asn Asp Leu Arg Pro Glu Ala Arg Pro Gly Gly Trp
290 295 300

Val Gly Glu Trp Leu Ala Gln Tyr Pro Asp Asn Tyr Glu Val Val Asn
305 310 315 320

Thr Leu Ala Arg Gln Ile Lys Asp Ile Trp Lys Asn Asn Gln His His
325 330 335

Lys Asp Gly Gly Glu Pro Tyr Lys Leu Ala Gln Arg Leu Ala Met Leu
340 345 350

Ala His Glu Ile Asp Ala Val Pro Ala Trp Asn Cys Lys Ser Gly Lys
355 360 365

Asp Arg Thr Gly Met Met Asp Ser Glu Ile Lys Arg Glu Ile Ile Ser
370 375 380

Leu His Gln Thr His Met Leu Ser Ala Pro Gly Ser Leu Pro Asp Ser
385 390 395 400

Gly Gly Gln Lys Ile Phe Gln Lys Val Leu Leu Asn Ser Gly Asn Leu
405 410 415

Glu Ile Gln Lys Gln Asn Thr Gly Gly Ala Gly Asn Lys Val Met Lys
420 425 430

Asn

<210> 10

<211> 538

<212> PRT

<213> CAC05808 IpgD, secreted by the Mxi-Spa machinery, modulates entry of bacteria into epithelial cells [Shigella flexneri]

<400> 10

Met His Ile Thr Asn Leu Gly Leu His Gln Val Ser Phe Gln Ser Gly
1 5 10 15

Asp Ser Tyr Lys Gly Ala Glu Glu Thr Gly Lys His Lys Gly Val Ser
20 25 30

Val Ile Ser Tyr Gln Arg Val Lys Asn Gly Glu Arg Asn Lys Gly Ile
35 40 45

Glu Ala Leu Asn Arg Leu Tyr Leu Gln Asn Gln Thr Ser Leu Thr Gly
50 55 60

Lys Ser Leu Leu Phe Ala Arg Asp Lys Ala Glu Val Phe Cys Glu Ala
65 70 75 80

Ile Lys Leu Ala Gly Gly Asp Thr Ser Lys Ile Lys Ala Met Met Glu
85 90 95

Arg Leu Asp Thr Tyr Lys Leu Gly Glu Val Asn Lys Arg His Ile Asn
100 105 110

Glu Leu Asn Lys Val Ile Ser Glu Glu Ile Arg Ala Gln Leu Gly Ile
115 120 125

Lys Asn Lys Lys Glu Leu Gln Thr Lys Ile Lys Gln Ile Phe Thr Asp
130 135 140

Tyr Leu Asn Asn Lys Asn Trp Gly Pro Val Asn Lys Asn Ile Ser His
145 150 155 160

His Gly Lys Asn Tyr Ser Phe Gln Leu Thr Pro Ala Ser His Met Lys
165 170 175

Ile Gly Asn Lys Asn Ile Phe Val Lys Glu Tyr Asn Gly Lys Gly Ile
180 185 190

Cys Cys Ala Ser Thr Arg Glu Arg Asp His Ile Ala Asn Met Trp Leu
195 200 205

Ser Lys Val Val Asp Asp Glu Gly Lys Glu Ile Phe Ser Gly Ile Arg

```
    210                 215                 220

His Gly Val Ile Ser Ala Tyr Gly Leu Lys Lys Asn Ser Ser Glu Arg
225                 230                 235                 240

Ala Val Ala Ala Arg Asn Lys Ala Glu Glu Leu Val Ser Ala Ala Leu
                245                 250                 255

Tyr Ser Arg Pro Glu Leu Leu Ser Gln Ala Leu Ser Gly Lys Thr Val
            260                 265                 270

Asp Leu Lys Ile Val Ser Thr Ser Leu Leu Thr Pro Thr Ser Leu Thr
        275                 280                 285

Gly Gly Glu Glu Ser Met Leu Lys Asp Gln Val Ser Ala Leu Lys Gly
    290                 295                 300

Leu Asn Ser Lys Arg Gly Gly Pro Thr Lys Leu Leu Ile Arg Asn Ser
305                 310                 315                 320

Asp Gly Leu Leu Lys Glu Val Ser Val Asn Leu Lys Val Val Thr Phe
                325                 330                 335

Asn Phe Gly Val Asn Glu Leu Ala Leu Lys Met Gly Leu Gly Trp Arg
            340                 345                 350

Asn Val Asp Lys Leu Asn Asp Glu Ser Ile Cys Ser Leu Leu Gly Asp
        355                 360                 365

Asn Phe Leu Lys Asn Gly Val Ile Gly Gly Trp Ala Ala Glu Ala Ile
    370                 375                 380

Glu Lys Asn Pro Pro Cys Lys Asn Asp Val Ile Tyr Leu Ala Asn Gln
385                 390                 395                 400

Ile Lys Glu Ile Val Asn Asn Lys Leu Gln Lys Asn Asp Asn Gly Glu
                405                 410                 415

Pro Tyr Lys Leu Ser Gln Arg Val Thr Leu Leu Ala Tyr Thr Ile Gly
            420                 425                 430

Ala Val Pro Cys Trp Asn Cys Lys Ser Gly Lys Asp Arg Thr Gly Met
        435                 440                 445

Gln Asp Ala Glu Ile Lys Arg Glu Ile Ile Arg Lys His Glu Thr Gly
    450                 455                 460

Gln Phe Ser Gln Leu Asn Ser Lys Leu Ser Ser Glu Glu Lys Arg Leu
465                 470                 475                 480
```

Phe Ser Thr Ile Leu Met Asn Ser Gly Asn Met Glu Ile Gln Glu Met
485 490 495

Asn Thr Gly Val Pro Gly Asn Lys Val Met Lys Lys Leu Pro Leu Ser
500 505 510

Ser Leu Glu Leu Ser Tyr Ser Glu Arg Ile Gly Asp Pro Lys Ile Trp
515 520 525

Asn Met Val Lys Gly Tyr Ser Ser Phe Val
530 535

<210> 11

<211> 288

<212> PRT

<213> AAC69766 targeted effector protein [Yersinia pestis] YopJ

<400> 11

Met Ile Gly Pro Ile Ser Gln Ile Asn Ile Ser Gly Gly Leu Ser Glu
1 5 10 15

Lys Glu Thr Ser Ser Leu Ile Ser Asn Glu Glu Leu Lys Asn Ile Ile
20 25 30

Thr Gln Leu Glu Thr Asp Ile Ser Asp Gly Ser Trp Phe His Lys Asn
35 40 45

Tyr Ser Arg Met Asp Val Glu Val Met Pro Ala Leu Val Ile Gln Ala
50 55 60

Asn Asn Lys Tyr Pro Glu Met Asn Leu Asn Leu Val Thr Ser Pro Leu
65 70 75 80

Asp Leu Ser Ile Glu Ile Lys Asn Val Ile Glu Asn Gly Val Arg Ser
85 90 95

Ser Arg Phe Ile Ile Asn Met Gly Glu Gly Gly Ile His Phe Ser Val
100 105 110

Ile Asp Tyr Lys His Ile Asn Gly Lys Thr Ser Leu Ile Leu Phe Glu
115 120 125

Pro Ala Asn Phe Asn Ser Met Gly Pro Ala Met Leu Ala Ile Arg Thr
130 135 140

Lys Thr Ala Ile Glu Arg Tyr Gln Leu Pro Asp Cys His Phe Ser Met
145 150 155 160

Val Glu Met Asp Ile Gln Arg Ser Ser Ser Glu Cys Gly Ile Phe Ser
165 170 175

Leu Ala Leu Ala Lys Lys Leu Tyr Ile Glu Arg Asp Ser Leu Leu Lys
180 185 190

Ile His Glu Asp Asn Ile Lys Gly Ile Leu Ser Asp Gly Glu Asn Pro
195 200 205

Leu Pro His Asp Lys Leu Asp Pro Tyr Leu Pro Val Thr Phe Tyr Lys
210 215 220

His Thr Gln Gly Lys Lys Arg Leu Asn Glu Tyr Leu Asn Thr Asn Pro
225 230 235 240

Gln Gly Val Gly Thr Val Val Asn Lys Lys Asn Glu Thr Ile Val Asn
245 250 255

Arg Phe Asp Asn Asn Lys Ser Ile Val Asp Gly Lys Glu Leu Ser Val
260 265 270

Ser Val His Lys Lys Arg Ile Ala Glu Tyr Lys Thr Leu Leu Lys Val
275 280 285

<210> 12

<211> 180

<212> PRT

<213> AAC02071 SopE [Salmonella typhimurium]

<400> 12

Met Thr Lys Ile Thr Leu Ser Pro Gln Asn Phe Arg Ile Gln Lys Gln
1 5 10 15

Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu Ala
20 25 30

Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser Lys
35 40 45

Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala Thr
50 55 60

His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu Thr Asn
65 70 75 80

Lys Val Val Lys Asp Phe Met Leu Gln Thr Leu Asn Asp Ile Asp Ile

85 90 95

Arg Gly Ser Ala Ser Lys Asp Pro Ala Tyr Ala Ser Gln Thr Arg Glu
100 105 110

Ala Ile Leu Ser Ala Val Tyr Ser Lys Asn Lys Asp Gln Cys Cys Asn
115 120 125

Leu Leu Ile Ser Lys Gly Ile Asn Ile Ala Pro Phe Leu Gln Glu Ile
130 135 140

Gly Glu Ala Ala Lys Asn Ala Gly Leu Pro Gly Thr Thr Lys Asn Asp
145 150 155 160

Val Phe Thr Pro Ser Gly Ala Gly Ala Asn Pro Phe Ile Thr Pro Leu
165 170 175

Ile Ser Ser Ala
180

<210> 13

<211> 543

<212> PRT

<213> AAC44349 protein tyrosine phosphatase SptP [Salmonella typhimurium]

<400> 13

Met Leu Lys Tyr Glu Glu Arg Lys Leu Asn Asn Leu Thr Leu Ser Ser
1 5 10 15

Phe Ser Lys Val Gly Val Ser Asn Asp Ala Arg Leu Tyr Ile Ala Lys
20 25 30

Glu Asn Thr Asp Lys Ala Tyr Val Ala Pro Glu Lys Phe Ser Ser Lys
35 40 45

Val Leu Thr Trp Leu Gly Lys Met Pro Leu Phe Lys Asn Thr Glu Val
50 55 60

Val Gln Lys His Thr Glu Asn Ile Arg Val Gln Asp Gln Lys Ile Leu
65 70 75 80

Gln Thr Phe Leu His Ala Leu Thr Glu Lys Tyr Gly Glu Thr Ala Val
85 90 95

Asn Asp Ala Leu Leu Met Ser Arg Ile Asn Met Asn Lys Pro Leu Thr
100 105 110

```
Gln Arg Leu Ala Val Gln Ile Thr Glu Cys Val Lys Ala Ala Asp Glu
        115                 120                 125

Gly Phe Ile Asn Leu Ile Lys Ser Lys Asp Asn Val Gly Val Arg Asn
    130                 135                 140

Ala Ala Leu Val Ile Lys Gly Gly Asp Thr Lys Val Ala Glu Lys Asn
145                 150                 155                 160

Asn Asp Val Gly Ala Glu Ser Lys Gln Pro Leu Leu Asp Ile Ala Leu
                165                 170                 175

Lys Gly Leu Lys Arg Thr Leu Pro Gln Leu Glu Gln Met Asp Gly Asn
            180                 185                 190

Ser Leu Arg Glu Asn Phe Gln Glu Met Ala Ser Gly Asn Gly Pro Leu
        195                 200                 205

Arg Ser Leu Met Thr Asn Leu Gln Asn Leu Asn Lys Ile Pro Glu Ala
    210                 215                 220

Lys Gln Leu Asn Asp Tyr Val Thr Thr Leu Thr Asn Ile Gln Val Gly
225                 230                 235                 240

Val Ala Arg Phe Ser Gln Trp Gly Thr Cys Gly Gly Glu Val Glu Arg
                245                 250                 255

Trp Val Asp Lys Ala Ser Thr His Glu Leu Thr Gln Ala Val Lys Lys
            260                 265                 270

Ile His Val Ile Ala Lys Glu Leu Lys Asn Val Thr Ala Glu Leu Glu
        275                 280                 285

Lys Ile Glu Ala Gly Ala Pro Met Pro Gln Thr Met Ser Gly Pro Thr
    290                 295                 300

Leu Gly Leu Ala Arg Phe Ala Val Ser Ser Ile Pro Ile Asn Gln Gln
305                 310                 315                 320

Thr Gln Val Lys Leu Ser Asp Gly Met Pro Val Pro Val Asn Thr Leu
                325                 330                 335

Thr Phe Asp Gly Lys Pro Val Ala Leu Ala Gly Ser Tyr Pro Lys Asn
            340                 345                 350

Thr Pro Asp Ala Leu Glu Ala His Met Lys Met Leu Leu Glu Lys Glu
        355                 360                 365

Cys Ser Cys Leu Val Val Leu Thr Ser Glu Asp Gln Met Gln Ala Lys
```

370 375 380

Gln Leu Pro Pro Tyr Phe Arg Gly Ser Tyr Thr Phe Gly Glu Val His
385 390 395 400

Thr Asn Ser Gln Lys Val Ser Ser Ala Ser Gln Gly Glu Ala Ile Asp
405 410 415

Gln Tyr Asn Met Gln Leu Ser Cys Gly Glu Lys Arg Tyr Thr Ile Pro
420 425 430

Val Leu His Val Lys Asn Trp Pro Asp His Gln Pro Leu Pro Ser Thr
435 440 445

Asp Gln Leu Glu Tyr Leu Ala Asp Arg Val Lys Asn Ser Asn Gln Asn
450 455 460

Gly Ala Pro Gly Arg Ser Ser Ser Asp Lys His Leu Pro Met Ile His
465 470 475 480

Cys Leu Gly Gly Val Gly Arg Thr Gly Thr Met Ala Ala Ala Leu Val
485 490 495

Leu Lys Asp Asn Pro His Ser Asn Leu Glu Gln Val Arg Ala Asp Phe
500 505 510

Arg Asp Ser Arg Asn Asn Arg Met Leu Glu Asp Ala Ser Gln Phe Val
515 520 525

Gln Leu Lys Ala Met Gln Ala Gln Leu Leu Met Thr Thr Ala Ser
530 535 540

<210> 14

<211> 219

<212> PRT

<213> NP_047628 targeted effector [Yersinia pestis] YopE

<400> 14

Met Lys Ile Ser Ser Phe Ile Ser Thr Ser Leu Pro Leu Pro Thr Ser
1 5 10 15

Val Ser Gly Ser Ser Ser Val Gly Glu Met Ser Gly Arg Ser Val Ser
20 25 30

Gln Gln Thr Ser Asp Gln Tyr Ala Asn Asn Leu Ala Gly Arg Thr Glu
35 40 45

Ser Pro Gln Gly Ser Ser Leu Ala Ser Arg Ile Ile Glu Arg Leu Ser
50 55 60

Ser Val Ala His Ser Val Ile Gly Phe Ile Gln Arg Met Phe Ser Glu
65 70 75 80

Gly Ser His Lys Pro Val Val Thr Pro Ala Pro Thr Pro Ala Gln Met
85 90 95

Pro Ser Pro Thr Ser Phe Ser Asp Ser Ile Lys Gln Leu Ala Ala Glu
100 105 110

Thr Leu Pro Lys Tyr Met Gln Gln Leu Asn Ser Leu Asp Ala Glu Met
115 120 125

Leu Gln Lys Asn His Asp Gln Phe Ala Thr Gly Ser Gly Pro Leu Arg
130 135 140

Gly Ser Ile Thr Gln Cys Gln Gly Leu Met Gln Phe Cys Gly Gly Glu
145 150 155 160

Leu Gln Ala Glu Ala Ser Ala Ile Leu Asn Thr Pro Val Cys Gly Ile
165 170 175

Pro Phe Ser Gln Trp Gly Thr Ile Gly Gly Ala Ala Ser Ala Tyr Val
180 185 190

Ala Ser Gly Val Asp Leu Thr Gln Ala Ala Asn Glu Ile Lys Gly Leu
195 200 205

Ala Gln Gln Met Gln Lys Leu Leu Ser Leu Met
210 215

<210> 15

<211> 453

<212> PRT

<213> AAK39624 exoenzyme S [Pseudomonas aeruginosa]

<400> 15

Met His Ile Gln Ser Leu Gln Gln Ser Pro Ser Phe Ala Val Glu Leu
1 5 10 15

His Gln Ala Ala Ser Gly Arg Leu Gly Gln Ile Glu Ala Arg Gln Val
20 25 30

Ala Thr Pro Ser Glu Ala Gln Gln Leu Ala Gln Arg Gln Asp Ala Pro
35 40 45

```
Lys Gly Glu Gly Leu Leu Ala Arg Leu Gly Ala Ala Leu Val Arg Pro
 50                  55                  60

Phe Val Ala Ile Met Asp Trp Leu Gly Lys Leu Leu Gly Ser His Ala
65                  70                  75                  80

Arg Thr Gly Pro Gln Pro Ser Gln Asp Ala Gln Pro Ala Val Met Ser
                85                  90                  95

Ser Ala Val Val Phe Lys Gln Met Val Leu Gln Gln Ala Leu Pro Met
            100                 105                 110

Thr Leu Lys Gly Leu Asp Lys Ala Ser Glu Leu Ala Thr Leu Thr Pro
        115                 120                 125

Glu Gly Leu Ala Arg Glu His Ser Arg Leu Ala Ser Gly Asp Gly Ala
    130                 135                 140

Leu Arg Ser Leu Ser Thr Ala Leu Ala Gly Ile Arg Ala Gly Ser Gln
145                 150                 155                 160

Val Glu Glu Ser Arg Ile Gln Ala Gly Arg Leu Leu Glu Arg Ser Ile
                165                 170                 175

Gly Gly Ile Ala Leu Gln Gln Trp Gly Thr Thr Gly Gly Ala Ala Ser
            180                 185                 190

Gln Leu Val Leu Asp Ala Ser Pro Glu Leu Arg Arg Glu Ile Thr Asp
        195                 200                 205

Gln Leu His Gln Val Met Ser Glu Val Ala Leu Leu Arg Gln Ala Val
    210                 215                 220

Glu Ser Glu Val Ser Arg Val Ser Ala Asp Lys Ala Leu Ala Asp Gly
225                 230                 235                 240

Leu Val Lys Arg Phe Gly Ala Asp Ala Glu Lys Tyr Leu Gly Arg Gln
                245                 250                 255

Pro Gly Gly Ile His Ser Asp Ala Glu Val Met Ala Leu Gly Leu Tyr
            260                 265                 270

Thr Gly Ile His Tyr Ala Asp Leu Asn Arg Ala Leu Arg Gln Gly Gln
        275                 280                 285

Glu Leu Asp Ala Gly Gln Lys Leu Ile Asp Gln Gly Met Ser Ala Ala
    290                 295                 300

Phe Glu Lys Ser Gly Gln Ala Glu Gln Val Val Lys Thr Phe Arg Gly
```

305 310 315 320

Thr Arg Gly Gly Asp Ala Phe Asn Ala Val Glu Glu Gly Lys Val Gly
325 330 335

His Asp Asp Gly Tyr Leu Ser Thr Ser Leu Asn Pro Gly Val Ala Arg
340 345 350

Ser Phe Gly Gln Gly Thr Ile Ser Thr Val Phe Gly Arg Ser Gly Ile
355 360 365

Asp Val Ser Gly Ile Ser Asn Tyr Lys Asn Glu Lys Glu Ile Leu Tyr
370 375 380

Asn Lys Glu Thr Asp Met Arg Val Leu Leu Ser Ala Ser Asp Glu Gln
385 390 395 400

Gly Val Thr Arg Arg Val Leu Glu Glu Ala Ala Leu Gly Glu Gln Ser
405 410 415

Gly His Ser Gln Gly Leu Leu Asp Ala Leu Asp Leu Ala Ser Lys Pro
420 425 430

Glu Arg Ser Gly Glu Val Gln Glu Gln Asp Val Arg Leu Arg Met Arg
435 440 445

Gly Leu Asp Leu Ala
450

<210> 16

<211> 457

<212> PRT

<213> AAG03434 exoenzyme T [Pseudomonas aeruginosa]

<400> 16

Met His Ile Gln Ser Ser Gln Gln Asn Pro Ser Phe Val Ala Glu Leu
1 5 10 15

Ser Gln Ala Val Ala Gly Arg Leu Gly Gln Val Glu Ala Arg Gln Val
20 25 30

Ala Thr Pro Arg Glu Ala Gln Gln Leu Ala Gln Arg Gln Glu Ala Pro
35 40 45

Lys Gly Glu Gly Leu Leu Ser Arg Leu Gly Ala Ala Leu Ala Arg Pro
50 55 60

```
Phe Val Ala Ile Ile Glu Trp Leu Gly Lys Leu Leu Gly Ser Arg Ala
65                  70                  75                  80

His Ala Ala Thr Gln Ala Pro Leu Ser Arg Gln Asp Ala Pro Pro Ala
                85                  90                  95

Ala Ser Leu Ser Ala Ala Glu Ile Lys Gln Met Met Leu Gln Lys Ala
            100                 105                 110

Leu Pro Leu Thr Leu Gly Gly Leu Gly Lys Ala Ser Glu Leu Ala Thr
        115                 120                 125

Leu Thr Ala Glu Arg Leu Ala Lys Asp His Thr Arg Leu Ala Ser Gly
    130                 135                 140

Asp Gly Ala Leu Arg Ser Leu Ala Thr Ala Leu Val Gly Ile Arg Asp
145                 150                 155                 160

Gly Ser Arg Ile Glu Ala Ser Arg Thr Gln Ala Ala Arg Leu Leu Glu
                165                 170                 175

Gln Ser Val Gly Gly Ile Ala Leu Gln Gln Trp Gly Thr Ala Gly Gly
            180                 185                 190

Ala Ala Ser Gln His Val Leu Ser Ala Ser Pro Glu Gln Leu Arg Glu
        195                 200                 205

Ile Ala Val Gln Leu His Ala Val Met Asp Lys Val Ala Leu Leu Arg
    210                 215                 220

His Ala Val Glu Ser Glu Val Lys Gly Glu Pro Val Asp Lys Ala Leu
225                 230                 235                 240

Ala Asp Gly Leu Val Glu His Phe Gly Leu Glu Ala Glu Gln Tyr Leu
                245                 250                 255

Gly Glu His Pro Asp Gly Pro Tyr Ser Asp Ala Glu Val Met Ala Leu
            260                 265                 270

Gly Leu Tyr Thr Asn Gly Glu Tyr Gln His Leu Asn Arg Ser Leu Arg
        275                 280                 285

Gln Gly Arg Glu Leu Asp Ala Gly Gln Ala Leu Ile Asp Arg Gly Met
    290                 295                 300

Ser Ala Ala Phe Glu Lys Ser Gly Pro Ala Glu Gln Val Val Lys Thr
305                 310                 315                 320

Phe Arg Gly Thr Gln Gly Arg Asp Ala Phe Glu Ala Val Lys Glu Gly
                325                 330                 335
```

Gln Val Gly His Asp Ala Gly Tyr Leu Ser Thr Ser Arg Asp Pro Gly
340 345 350

Val Ala Arg Ser Phe Ala Gly Gln Gly Thr Ile Thr Thr Leu Phe Gly
355 360 365

Arg Ser Gly Ile Asp Val Ser Glu Ile Ser Ile Glu Gly Asp Glu Gln
370 375 380

Glu Ile Leu Tyr Asp Lys Gly Thr Asp Met Arg Val Leu Leu Ser Ala
385 390 395 400

Lys Asp Gly Gln Gly Val Thr Arg Arg Val Leu Glu Glu Ala Thr Leu
405 410 415

Gly Glu Arg Ser Gly His Gly Glu Gly Leu Leu Asp Ala Leu Asp Leu
420 425 430

Ala Thr Gly Thr Asp Arg Ser Gly Lys Pro Gln Glu Gln Asp Leu Arg
435 440 445

Leu Arg Met Arg Gly Leu Asp Leu Ala
450 455

<210> 17

<211> 322

<212> PRT

<213> NP_047619 Yop targeted effector [Yersinia pestis] YopT

<400> 17

Met Asn Ser Ile His Gly His Tyr His Ile Gln Leu Ser Asn Tyr Ser
1 5 10 15

Ala Gly Glu Asn Leu Gln Ser Ala Thr Leu Thr Glu Gly Val Ile Gly
20 25 30

Ala His Arg Val Lys Val Glu Thr Ala Leu Ser His Ser Asn Leu Gln
35 40 45

Lys Lys Leu Ser Ala Thr Ile Lys His Asn Gln Ser Gly Arg Ser Met
50 55 60

Leu Asp Arg Lys Leu Thr Ser Asp Gly Lys Ala Asn Gln Arg Ser Ser
65 70 75 80

Phe Thr Phe Ser Met Ile Met Tyr Arg Met Ile His Phe Val Leu Ser

85 90 95

Thr Arg Val Pro Ala Val Arg Glu Ser Val Ala Asn Tyr Gly Gly Asn
100 105 110

Ile Asn Phe Lys Phe Ala Gln Thr Lys Gly Ala Phe Leu His Lys Ile
115 120 125

Ile Lys His Ser Asp Thr Ala Ser Gly Val Cys Glu Ala Leu Cys Ala
130 135 140

His Trp Ile Arg Ser His Ala Gln Gly Gln Ser Leu Phe Asp Gln Leu
145 150 155 160

Tyr Val Gly Gly Arg Lys Gly Lys Phe Gln Ile Asp Thr Leu Tyr Ser
165 170 175

Ile Lys Gln Leu Gln Ile Asp Gly Cys Lys Ala Asp Val Asp Gln Asp
180 185 190

Glu Val Thr Leu Asp Trp Phe Lys Lys Asn Gly Ile Ser Glu Arg Met
195 200 205

Ile Glu Arg His Cys Leu Leu Arg Pro Val Asp Val Thr Gly Thr Thr
210 215 220

Glu Ser Glu Gly Leu Asp Gln Leu Leu Asn Ala Ile Leu Asp Thr His
225 230 235 240

Gly Ile Gly Tyr Gly Tyr Lys Lys Ile His Leu Ser Gly Gln Met Ser
245 250 255

Ala His Ala Ile Ala Ala Tyr Val Asn Glu Lys Ser Gly Val Thr Phe
260 265 270

Phe Asp Pro Asn Phe Gly Glu Phe His Phe Ser Asp Lys Glu Lys Phe
275 280 285

Arg Lys Trp Phe Thr Asn Ser Phe Trp Gly Asn Ser Met Tyr His Tyr
290 295 300

Pro Leu Gly Val Gly Gln Arg Phe Arg Val Leu Thr Phe Asp Ser Lys
305 310 315 320

Glu Val

<210> 18

<211> 729

<212> PRT

<213> NP_052380 protein kinase YopO [Yersinia enterocolitica]

<400> 18

Met Lys Ile Met Gly Thr Met Pro Pro Ser Ile Ser Leu Ala Lys Ala
1 5 10 15

His Glu Arg Ile Ser Gln His Trp Gln Asn Pro Val Gly Glu Leu Asn
20 25 30

Ile Gly Gly Lys Arg Tyr Arg Ile Ile Asp Asn Gln Val Leu Arg Leu
35 40 45

Asn Pro His Ser Gly Phe Ser Leu Phe Arg Glu Gly Val Gly Lys Ile
50 55 60

Phe Ser Gly Lys Met Phe Asn Phe Ser Ile Ala Arg Asn Leu Thr Glu
65 70 75 80

Thr Leu His Ala Ala Gln Lys Thr Thr Ser Gln Glu Leu Arg Ser Asp
85 90 95

Ile Pro Asn Ala Leu Ser Asn Leu Phe Gly Ala Lys Pro Gln Thr Glu
100 105 110

Leu Pro Leu Gly Trp Lys Gly Lys Pro Leu Ser Gly Ala Pro Asp Leu
115 120 125

Glu Gly Met Arg Val Ala Glu Thr Asp Lys Phe Ala Glu Gly Glu Ser
130 135 140

His Ile Ser Ile Ile Glu Thr Lys Asp Asn Gln Arg Leu Val Ala Lys
145 150 155 160

Ile Glu Arg Ser Ile Ala Glu Gly His Leu Phe Ala Glu Leu Glu Ala
165 170 175

Tyr Lys His Ile Tyr Lys Thr Ala Gly Lys His Pro Asn Leu Ala Asn
180 185 190

Val His Gly Met Ala Val Val Pro Tyr Gly Asn Arg Lys Glu Glu Ala
195 200 205

Leu Leu Met Asp Glu Val Asp Gly Trp Arg Cys Ser Asp Thr Leu Arg
210 215 220

Ser Leu Ala Asp Ser Trp Lys Gln Gly Lys Ile Asn Ser Glu Ala Tyr
225 230 235 240

Trp Gly Thr Ile Lys Phe Ile Ala His Arg Leu Leu Asp Val Thr Asn
245 250 255

His Leu Ala Lys Ala Gly Ile Val His Asn Asp Ile Lys Pro Gly Asn
260 265 270

Val Val Phe Asp Arg Ala Ser Gly Glu Pro Val Val Ile Asp Leu Gly
275 280 285

Leu His Ser Arg Ser Gly Glu Gln Pro Lys Gly Phe Thr Glu Ser Phe
290 295 300

Lys Ala Pro Glu Leu Gly Val Gly Asn Leu Gly Ala Ser Glu Lys Ser
305 310 315 320

Asp Val Phe Leu Val Val Ser Thr Leu Leu His Gly Ile Glu Gly Phe
325 330 335

Glu Lys Asp Pro Glu Ile Lys Pro Asn Gln Gly Leu Arg Phe Ile Thr
340 345 350

Ser Glu Pro Ala His Val Met Asp Glu Asn Gly Tyr Pro Ile His Arg
355 360 365

Pro Gly Ile Ala Gly Val Glu Thr Ala Tyr Thr Arg Phe Ile Thr Asp
370 375 380

Ile Leu Gly Val Ser Ala Asp Ser Arg Pro Asp Ser Asn Glu Ala Arg
385 390 395 400

Leu His Glu Phe Leu Ser Asp Gly Thr Ile Asp Glu Glu Ser Ala Lys
405 410 415

Gln Ile Leu Lys Asp Thr Leu Thr Gly Glu Met Ser Pro Leu Ser Thr
420 425 430

Asp Val Arg Arg Ile Thr Pro Lys Lys Leu Arg Glu Leu Ser Asp Leu
435 440 445

Leu Arg Thr His Leu Ser Ser Ala Ala Thr Lys Gln Leu Asp Met Gly
450 455 460

Val Val Leu Ser Asp Leu Asp Thr Met Leu Val Thr Leu Asp Lys Ala
465 470 475 480

Glu Arg Glu Gly Gly Val Asp Lys Asp Gln Leu Lys Ser Phe Asn Ser
485 490 495

Leu Ile Leu Lys Thr Tyr Ser Val Ile Glu Asp Tyr Val Lys Gly Arg

500 505 510

Glu Gly Asp Thr Lys Ser Ser Ser Ala Glu Val Ser Pro Tyr His Arg
515 520 525

Ser Asn Phe Met Leu Ser Ile Ala Glu Pro Ser Leu Gln Arg Ile Gln
530 535 540

Lys His Leu Asp Gln Thr His Ser Phe Ser Asp Ile Gly Ser Leu Val
545 550 555 560

Arg Ala His Lys His Leu Glu Thr Leu Leu Glu Val Leu Val Thr Leu
565 570 575

Ser Pro Gln Gly Gln Pro Val Ser Ser Glu Thr Tyr Ser Phe Leu Asn
580 585 590

Arg Leu Ala Glu Ala Lys Val Thr Leu Ser Gln Gln Leu Asp Thr Leu
595 600 605

Gln Gln Gln Gln Glu Ser Ala Lys Ala Gln Leu Ser Ile Leu Ile Asn
610 615 620

Arg Ser Gly Ser Trp Ala Asp Val Ala Arg Gln Ser Leu Gln Arg Phe
625 630 635 640

Asp Ser Thr Arg Pro Val Val Lys Phe Gly Thr Glu Gln Tyr Thr Ala
645 650 655

Ile His Arg Gln Met Met Ala Ala His Ala Ala Ile Thr Leu Gln Glu
660 665 670

Val Ser Glu Phe Thr Asp Asp Met Arg Asn Phe Thr Ala Asp Ser Ile
675 680 685

Pro Leu Leu Ile Arg Leu Gly Arg Ser Ser Leu Ile Asp Glu His Leu
690 695 700

Val Glu Gln Arg Glu Lys Leu Arg Glu Leu Thr Thr Ile Ala Glu Arg
705 710 715 720

Leu Asn Arg Leu Glu Arg Glu Trp Met
725

<210> 19

<211> 129

<212> PRT

<213> AAF82095 outer protein AvrA [Salmonella enterica subsp. enterica

serovar Dublin]

<400> 19

Val Met Asp Gly Lys Thr Ser Val Ile Leu Phe Glu Pro Ala Ala Cys
1 5 10 15

Ser Ala Phe Gly Pro Ala Leu Leu Ala Leu Arg Thr Lys Ala Ala Leu
20 25 30

Glu Arg Glu Gln Leu Pro Asp Cys Tyr Phe Ala Met Val Glu Leu Asp
35 40 45

Ile Gln Arg Ser Ser Ser Glu Cys Gly Ile Phe Ser Leu Ala Leu Ala
50 55 60

Lys Lys Leu Gln Leu Glu Phe Met Asn Leu Val Lys Ile His Glu Asp
65 70 75 80

Asn Ile Cys Glu Arg Leu Cys Gly Glu Glu Pro Phe Leu Pro Ser Asp
85 90 95

Lys Ala Asp Arg Tyr Leu Pro Val Ser Phe Tyr Lys His Thr Gln Gly
100 105 110

Val Gln Arg Leu Asn Glu Tyr Val Glu Ala Asn Pro Ala Ala Gly Ser
115 120 125

Ser

<210> 20

<211> 133

<212> PRT

<213> AAC44300 SpiC [Salmonella typhimurium]

<400> 20

Met Ser Glu Glu Gly Phe Met Leu Ala Val Leu Lys Gly Ile Pro Leu
1 5 10 15

Ile Gln Asp Ile Arg Ala Glu Gly Asn Ser Arg Ser Trp Ile Met Thr
20 25 30

Ile Asp Gly His Pro Ala Arg Gly Glu Ile Phe Ser Glu Ala Phe Ser
35 40 45

Ile Ser Leu Phe Leu Asn Asp Leu Glu Ser Leu Pro Lys Pro Cys Leu
50 55 60

Ala Tyr Val Thr Leu Leu Leu Ala Ala His Pro Asp Val His Asp Tyr
65 70 75 80

Ala Ile Gln Leu Thr Ala Asp Gly Gly Trp Leu Asn Gly Tyr Tyr Thr
85 90 95

Thr Ser Ser Ser Ser Glu Leu Ile Ala Ile Glu Ile Glu Lys His Leu
100 105 110

Ala Leu Thr Cys Ile Leu Lys Asn Val Ile Arg Asn His His Lys Leu
115 120 125

Tyr Ser Gly Gly Val
130

<210> 21

<211> 1212

<212> PRT

<213> Protein sequence for SigD with the first 29 codons removed, thrombin linker, diphtheria translocation domain, TeNT-HC

<400> 21

Met Gln Ile Leu Ser Gly Gln Gly Lys Ala Pro Ala Lys Ala Pro Asp
1 5 10 15

Ala Arg Pro Glu Ile Ile Val Leu Arg Glu Pro Gly Ala Thr Trp Gly
20 25 30

Asn Tyr Leu Gln His Gln Lys Ala Ser Asn His Ser Leu His Asn Leu
35 40 45

Tyr Asn Leu Gln Arg Asp Leu Leu Thr Val Ala Ala Thr Val Leu Gly
50 55 60

Lys Gln Asp Pro Val Leu Thr Ser Met Ala Asn Gln Met Glu Leu Ala
65 70 75 80

Lys Val Lys Ala Asp Arg Pro Ala Thr Lys Gln Glu Glu Ala Ala Ala
85 90 95

Lys Ala Leu Lys Lys Asn Leu Ile Glu Leu Ile Ala Ala Arg Thr Gln
100 105 110

Gln Gln Asp Gly Leu Pro Ala Lys Glu Ala His Arg Phe Ala Ala Val

```
        115                 120                 125

Ala Phe Arg Asp Ala Gln Val Lys Gln Leu Asn Asn Gln Pro Trp Gln
    130                 135                 140

Thr Ile Lys Asn Thr Leu Thr His Asn Gly His His Tyr Thr Asn Thr
145                 150                 155                 160

Gln Leu Pro Ala Ala Glu Met Lys Ile Gly Ala Lys Asp Ile Phe Pro
                165                 170                 175

Ser Ala Tyr Glu Gly Lys Gly Val Cys Ser Trp Asp Thr Lys Asn Ile
            180                 185                 190

His His Ala Asn Asn Leu Trp Met Ser Thr Val Ser Val His Glu Asp
        195                 200                 205

Gly Lys Asp Lys Thr Leu Phe Phe Asp Gly Ile Arg His Gly Val Leu
    210                 215                 220

Ser Pro Tyr His Glu Lys Asp Pro Leu Leu Arg His Val Gly Ala Glu
225                 230                 235                 240

Asn Lys Ala Lys Glu Val Leu Thr Ala Ala Leu Phe Ser Lys Pro Glu
                245                 250                 255

Leu Leu Asn Lys Ala Leu Ala Gly Glu Ala Val Ser Leu Lys Leu Val
            260                 265                 270

Ser Val Gly Leu Leu Thr Ala Ser Asn Ile Phe Gly Lys Glu Gly Thr
        275                 280                 285

Met Val Glu Asp Gln Met Arg Ala Trp Gln Ser Leu Thr Gln Pro Gly
    290                 295                 300

Lys Met Ile His Leu Lys Ile Arg Asn Lys Asp Gly Asp Leu Gln Thr
305                 310                 315                 320

Val Lys Ile Lys Pro Asp Val Val Ala Ala Phe Asn Val Gly Val Asn
                325                 330                 335

Glu Leu Ala Leu Lys Leu Gly Phe Gly Leu Lys Ala Ser Asp Ser Tyr
            340                 345                 350

Asn Ala Glu Ala Leu His Gln Leu Leu Gly Asn Asp Leu Arg Pro Glu
        355                 360                 365

Ala Arg Pro Gly Gly Trp Val Gly Glu Trp Leu Ala Gln Tyr Pro Asp
    370                 375                 380
```

```
Asn Tyr Glu Val Val Asn Thr Leu Ala Arg Gln Ile Lys Asp Ile Trp
385                 390                 395                 400

Lys Asn Asn Gln His His Lys Asp Gly Gly Glu Pro Tyr Lys Leu Ala
                405                 410                 415

Gln Arg Leu Ala Met Leu Ala His Glu Ile Asp Ala Val Pro Ala Trp
            420                 425                 430

Asn Cys Lys Ser Gly Lys Asp Arg Thr Gly Met Met Asp Ser Glu Ile
        435                 440                 445

Lys Gly Glu Ile Ile Ser Leu His Gln Thr His Met Leu Ser Ala Pro
    450                 455                 460

Gly Ser Leu Pro Asp Ser Gly Gly Gln Lys Ile Phe Gln Lys Val Leu
465                 470                 475                 480

Leu Asn Ser Gly Asn Leu Glu Ile Gln Lys Gln Asn Thr Gly Gly Ala
                485                 490                 495

Gly Asn Lys Val Met Lys Asn Leu Ser Pro Glu Val Leu Asn Leu Ser
            500                 505                 510

Tyr Gln Lys Arg Val Gly Asp Glu Asn Ile Trp Gln Ser Val Lys Gly
        515                 520                 525

Ile Ser Ser Leu Ile Thr Ser Arg Ser Cys Gly Leu Val Pro Arg Gly
    530                 535                 540

Ser Gly Pro Gly Ser Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu
545                 550                 555                 560

Asp Trp Asp Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu
                565                 570                 575

Lys Glu His Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys
            580                 585                 590

Thr Val Ser Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln
        595                 600                 605

Thr Ala Leu Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly
    610                 615                 620

Thr Asn Pro Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn
625                 630                 635                 640

Val Ala Gln Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr
                645                 650                 655
```

Thr Ala Ala Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile
660 665 670

Ala Asp Gly Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser
675 680 685

Ile Ala Leu Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly
690 695 700

Glu Leu Val Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile
705 710 715 720

Ile Asn Leu Phe Gln Val Val His Asn Ser Tyr Asn Arg Ser Ala Tyr
725 730 735

Ser Pro Gly His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val
740 745 750

Ser Trp Asn Thr Val Arg Ser Lys Asn Leu Asp Cys Trp Val Asp Asn
755 760 765

Glu Glu Asp Ile Asp Val Ile Leu Lys Lys Ser Thr Ile Leu Asn Leu
770 775 780

Asp Ile Asn Asn Asp Ile Ile Ser Asp Ile Ser Gly Phe Asn Ser Ser
785 790 795 800

Val Ile Thr Tyr Pro Asp Ala Gln Leu Val Pro Gly Ile Asn Gly Lys
805 810 815

Ala Ile His Leu Val Asn Asn Glu Ser Ser Glu Val Ile Val His Lys
820 825 830

Ala Met Asp Ile Glu Tyr Asn Asp Met Phe Asn Asn Phe Thr Val Ser
835 840 845

Phe Trp Leu Arg Val Pro Lys Val Ser Ala Ser His Leu Glu Gln Tyr
850 855 860

Gly Thr Asn Glu Tyr Ser Ile Ile Ser Ser Met Lys Lys His Ser Leu
865 870 875 880

Ser Ile Gly Ser Gly Trp Ser Val Ser Leu Lys Gly Asn Asn Leu Ile
885 890 895

Trp Thr Leu Lys Asp Ser Ala Gly Glu Val Arg Gln Ile Thr Phe Arg
900 905 910

Asp Leu Pro Asp Lys Phe Asn Ala Tyr Leu Ala Asn Lys Trp Val Phe

915 920 925

Ile Thr Ile Thr Asn Asp Arg Leu Ser Ser Ala Asn Leu Tyr Ile Asn
930 935 940

Gly Val Leu Met Gly Ser Ala Glu Ile Thr Gly Leu Gly Ala Ile Arg
945 950 955 960

Glu Asp Asn Asn Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn Asn Asn
965 970 975

Gln Tyr Val Ser Ile Asp Lys Phe Arg Ile Phe Cys Lys Ala Leu Asn
980 985 990

Pro Lys Glu Ile Glu Lys Leu Tyr Thr Ser Tyr Leu Ser Ile Thr Phe
995 1000 1005

Leu Arg Asp Phe Trp Gly Asn Pro Leu Arg Tyr Asp Thr Glu Tyr
1010 1015 1020

Tyr Leu Ile Pro Val Ala Ser Ser Ser Lys Asp Val Gln Leu Lys
1025 1030 1035

Asn Ile Thr Asp Tyr Met Tyr Leu Thr Asn Ala Pro Ser Tyr Thr
1040 1045 1050

Asn Gly Lys Leu Asn Ile Tyr Tyr Arg Arg Leu Tyr Asn Gly Leu
1055 1060 1065

Lys Phe Ile Ile Lys Arg Tyr Thr Pro Asn Asn Glu Ile Asp Ser
1070 1075 1080

Phe Val Lys Ser Gly Asp Phe Ile Lys Leu Tyr Val Ser Tyr Asn
1085 1090 1095

Asn Asn Glu His Ile Val Gly Tyr Pro Lys Asp Gly Asn Ala Phe
1100 1105 1110

Asn Asn Leu Asp Arg Ile Leu Arg Val Gly Tyr Asn Ala Pro Gly
1115 1120 1125

Ile Pro Leu Tyr Lys Lys Met Glu Ala Val Lys Leu Arg Asp Leu
1130 1135 1140

Lys Thr Tyr Ser Val Gln Leu Lys Leu Tyr Asp Asp Lys Asn Ala
1145 1150 1155

Ser Leu Gly Leu Val Gly Thr His Asn Gly Gln Ile Gly Asn Asp
1160 1165 1170

Pro Asn Arg Asp Ile Leu Ile Ala Ser Asn Trp Tyr Phe Asn His
1175 1180 1185

Leu Lys Asp Lys Ile Leu Gly Cys Asp Trp Tyr Phe Val Pro Thr
1190 1195 1200

Asp Glu Gly Trp Thr Asn Asp Leu Gln
1205 1210

<210> 22

<211> 1212

<212> PRT

<213> Protein sequence for SigD with the first 29 codons removed, factor Xa linker, diphtheria translocation domain, TeNT-HC

<400> 22

Met Gln Ile Leu Ser Gly Gln Gly Lys Ala Pro Ala Lys Ala Pro Asp
1 5 10 15

Ala Arg Pro Glu Ile Ile Val Leu Arg Glu Pro Gly Ala Thr Trp Gly
20 25 30

Asn Tyr Leu Gln His Gln Lys Ala Ser Asn His Ser Leu His Asn Leu
35 40 45

Tyr Asn Leu Gln Arg Asp Leu Leu Thr Val Ala Ala Thr Val Leu Gly
50 55 60

Lys Gln Asp Pro Val Leu Thr Ser Met Ala Asn Gln Met Glu Leu Ala
65 70 75 80

Lys Val Lys Ala Asp Arg Pro Ala Thr Lys Gln Glu Glu Ala Ala Ala
85 90 95

Lys Ala Leu Lys Lys Asn Leu Ile Glu Leu Ile Ala Ala Arg Thr Gln
100 105 110

Gln Gln Asp Gly Leu Pro Ala Lys Glu Ala His Arg Phe Ala Ala Val
115 120 125

Ala Phe Arg Asp Ala Gln Val Lys Gln Leu Asn Asn Gln Pro Trp Gln
130 135 140

Thr Ile Lys Asn Thr Leu Thr His Asn Gly His His Tyr Thr Asn Thr
145 150 155 160

Gln Leu Pro Ala Ala Glu Met Lys Ile Gly Ala Lys Asp Ile Phe Pro

165 170 175

Ser Ala Tyr Glu Gly Lys Gly Val Cys Ser Trp Asp Thr Lys Asn Ile
180 185 190

His His Ala Asn Asn Leu Trp Met Ser Thr Val Ser Val His Glu Asp
195 200 205

Gly Lys Asp Lys Thr Leu Phe Phe Asp Gly Ile Arg His Gly Val Leu
210 215 220

Ser Pro Tyr His Glu Lys Asp Pro Leu Leu Arg His Val Gly Ala Glu
225 230 235 240

Asn Lys Ala Lys Glu Val Leu Thr Ala Ala Leu Phe Ser Lys Pro Glu
245 250 255

Leu Leu Asn Lys Ala Leu Ala Gly Glu Ala Val Ser Leu Lys Leu Val
260 265 270

Ser Val Gly Leu Leu Thr Ala Ser Asn Ile Phe Gly Lys Glu Gly Thr
275 280 285

Met Val Glu Asp Gln Met Arg Ala Trp Gln Ser Leu Thr Gln Pro Gly
290 295 300

Lys Met Ile His Leu Lys Ile Arg Asn Lys Asp Gly Asp Leu Gln Thr
305 310 315 320

Val Lys Ile Lys Pro Asp Val Val Ala Ala Phe Asn Val Gly Val Asn
325 330 335

Glu Leu Ala Leu Lys Leu Gly Phe Gly Leu Lys Ala Ser Asp Ser Tyr
340 345 350

Asn Ala Glu Ala Leu His Gln Leu Leu Gly Asn Asp Leu Arg Pro Glu
355 360 365

Ala Arg Pro Gly Gly Trp Val Gly Glu Trp Leu Ala Gln Tyr Pro Asp
370 375 380

Asn Tyr Glu Val Val Asn Thr Leu Ala Arg Gln Ile Lys Asp Ile Trp
385 390 395 400

Lys Asn Asn Gln His His Lys Asp Gly Gly Glu Pro Tyr Lys Leu Ala
405 410 415

Gln Arg Leu Ala Met Leu Ala His Glu Ile Asp Ala Val Pro Ala Trp
420 425 430

```
Asn Cys Lys Ser Gly Lys Asp Arg Thr Gly Met Met Asp Ser Glu Ile
        435                 440                 445

Lys Gly Glu Ile Ile Ser Leu His Gln Thr His Met Leu Ser Ala Pro
    450                 455                 460

Gly Ser Leu Pro Asp Ser Gly Gly Gln Lys Ile Phe Gln Lys Val Leu
465                 470                 475                 480

Leu Asn Ser Gly Asn Leu Glu Ile Gln Lys Gln Asn Thr Gly Gly Ala
                485                 490                 495

Gly Asn Lys Val Met Lys Asn Leu Ser Pro Glu Val Leu Asn Leu Ser
            500                 505                 510

Tyr Gln Lys Arg Val Gly Asp Glu Asn Ile Trp Gln Ser Val Lys Gly
        515                 520                 525

Ile Ser Ser Leu Ile Thr Ser Arg Ser Cys Gly Ile Glu Gly Arg Ala
    530                 535                 540

Pro Gly Pro Gly Ser Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu
545                 550                 555                 560

Asp Trp Asp Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu
                565                 570                 575

Lys Glu His Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys
            580                 585                 590

Thr Val Ser Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln
        595                 600                 605

Thr Ala Leu Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly
    610                 615                 620

Thr Asn Pro Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn
625                 630                 635                 640

Val Ala Gln Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr
                645                 650                 655

Thr Ala Ala Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile
            660                 665                 670

Ala Asp Gly Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser
        675                 680                 685

Ile Ala Leu Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly
    690                 695                 700
```

Glu Leu Val Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile
705 710 715 720

Ile Asn Leu Phe Gln Val Val His Asn Ser Tyr Asn Arg Ser Ala Tyr
725 730 735

Ser Pro Gly His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val
740 745 750

Ser Trp Asn Thr Val Arg Ser Lys Asn Leu Asp Cys Trp Val Asp Asn
755 760 765

Glu Glu Asp Ile Asp Val Ile Leu Lys Lys Ser Thr Ile Leu Asn Leu
770 775 780

Asp Ile Asn Asn Asp Ile Ile Ser Asp Ile Ser Gly Phe Asn Ser Ser
785 790 795 800

Val Ile Thr Tyr Pro Asp Ala Gln Leu Val Pro Gly Ile Asn Gly Lys
805 810 815

Ala Ile His Leu Val Asn Asn Glu Ser Ser Glu Val Ile Val His Lys
820 825 830

Ala Met Asp Ile Glu Tyr Asn Asp Met Phe Asn Asn Phe Thr Val Ser
835 840 845

Phe Trp Leu Arg Val Pro Lys Val Ser Ala Ser His Leu Glu Gln Tyr
850 855 860

Gly Thr Asn Glu Tyr Ser Ile Ile Ser Ser Met Lys Lys His Ser Leu
865 870 875 880

Ser Ile Gly Ser Gly Trp Ser Val Ser Leu Lys Gly Asn Asn Leu Ile
885 890 895

Trp Thr Leu Lys Asp Ser Ala Gly Glu Val Arg Gln Ile Thr Phe Arg
900 905 910

Asp Leu Pro Asp Lys Phe Asn Ala Tyr Leu Ala Asn Lys Trp Val Phe
915 920 925

Ile Thr Ile Thr Asn Asp Arg Leu Ser Ser Ala Asn Leu Tyr Ile Asn
930 935 940

Gly Val Leu Met Gly Ser Ala Glu Ile Thr Gly Leu Gly Ala Ile Arg
945 950 955 960

Glu Asp Asn Asn Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn Asn Asn

965 970 975

Gln Tyr Val Ser Ile Asp Lys Phe Arg Ile Phe Cys Lys Ala Leu Asn
980 985 990

Pro Lys Glu Ile Glu Lys Leu Tyr Thr Ser Tyr Leu Ser Ile Thr Phe
995 1000 1005

Leu Arg Asp Phe Trp Gly Asn Pro Leu Arg Tyr Asp Thr Glu Tyr
1010 1015 1020

Tyr Leu Ile Pro Val Ala Ser Ser Ser Lys Asp Val Gln Leu Lys
1025 1030 1035

Asn Ile Thr Asp Tyr Met Tyr Leu Thr Asn Ala Pro Ser Tyr Thr
1040 1045 1050

Asn Gly Lys Leu Asn Ile Tyr Tyr Arg Arg Leu Tyr Asn Gly Leu
1055 1060 1065

Lys Phe Ile Ile Lys Arg Tyr Thr Pro Asn Asn Glu Ile Asp Ser
1070 1075 1080

Phe Val Lys Ser Gly Asp Phe Ile Lys Leu Tyr Val Ser Tyr Asn
1085 1090 1095

Asn Asn Glu His Ile Val Gly Tyr Pro Lys Asp Gly Asn Ala Phe
1100 1105 1110

Asn Asn Leu Asp Arg Ile Leu Arg Val Gly Tyr Asn Ala Pro Gly
1115 1120 1125

Ile Pro Leu Tyr Lys Lys Met Glu Ala Val Lys Leu Arg Asp Leu
1130 1135 1140

Lys Thr Tyr Ser Val Gln Leu Lys Leu Tyr Asp Asp Lys Asn Ala
1145 1150 1155

Ser Leu Gly Leu Val Gly Thr His Asn Gly Gln Ile Gly Asn Asp
1160 1165 1170

Pro Asn Arg Asp Ile Leu Ile Ala Ser Asn Trp Tyr Phe Asn His
1175 1180 1185

Leu Lys Asp Lys Ile Leu Gly Cys Asp Trp Tyr Phe Val Pro Thr
1190 1195 1200

Asp Glu Gly Trp Thr Asn Asp Leu Gln
1205 1210

<210> 23

<211> 1192

<212> PRT

<213> Protein sequence for SigD with the first 29 codons removed, thrombin linker, diphtheria toxin translocation domain, with BoNT/F-HC

<400> 23

```
Met Gln Ile Leu Ser Gly Gln Gly Lys Ala Pro Ala Lys Ala Pro Asp
1               5                  10                  15

Ala Arg Pro Glu Ile Ile Val Leu Arg Glu Pro Gly Ala Thr Trp Gly
            20                  25                  30

Asn Tyr Leu Gln His Gln Lys Ala Ser Asn His Ser Leu His Asn Leu
        35                  40                  45

Tyr Asn Leu Gln Arg Asp Leu Leu Thr Val Ala Ala Thr Val Leu Gly
    50                  55                  60

Lys Gln Asp Pro Val Leu Thr Ser Met Ala Asn Gln Met Glu Leu Ala
65                  70                  75                  80

Lys Val Lys Ala Asp Arg Pro Ala Thr Lys Gln Glu Glu Ala Ala Ala
                85                  90                  95

Lys Ala Leu Lys Lys Asn Leu Ile Glu Leu Ile Ala Ala Arg Thr Gln
            100                 105                 110

Gln Gln Asp Gly Leu Pro Ala Lys Glu Ala His Arg Phe Ala Ala Val
        115                 120                 125

Ala Phe Arg Asp Ala Gln Val Lys Gln Leu Asn Asn Gln Pro Trp Gln
    130                 135                 140

Thr Ile Lys Asn Thr Leu Thr His Asn Gly His His Tyr Thr Asn Thr
145                 150                 155                 160

Gln Leu Pro Ala Ala Glu Met Lys Ile Gly Ala Lys Asp Ile Phe Pro
                165                 170                 175

Ser Ala Tyr Glu Gly Lys Gly Val Cys Ser Trp Asp Thr Lys Asn Ile
            180                 185                 190

His His Ala Asn Asn Leu Trp Met Ser Thr Val Ser Val His Glu Asp
        195                 200                 205

Gly Lys Asp Lys Thr Leu Phe Phe Asp Gly Ile Arg His Gly Val Leu
```

210 215 220

Ser Pro Tyr His Glu Lys Asp Pro Leu Leu Arg His Val Gly Ala Glu
225 230 235 240

Asn Lys Ala Lys Glu Val Leu Thr Ala Ala Leu Phe Ser Lys Pro Glu
245 250 255

Leu Leu Asn Lys Ala Leu Ala Gly Glu Ala Val Ser Leu Lys Leu Val
260 265 270

Ser Val Gly Leu Leu Thr Ala Ser Asn Ile Phe Gly Lys Glu Gly Thr
275 280 285

Met Val Glu Asp Gln Met Arg Ala Trp Gln Ser Leu Thr Gln Pro Gly
290 295 300

Lys Met Ile His Leu Lys Ile Arg Asn Lys Asp Gly Asp Leu Gln Thr
305 310 315 320

Val Lys Ile Lys Pro Asp Val Val Ala Ala Phe Asn Val Gly Val Asn
325 330 335

Glu Leu Ala Leu Lys Leu Gly Phe Gly Leu Lys Ala Ser Asp Ser Tyr
340 345 350

Asn Ala Glu Ala Leu His Gln Leu Leu Gly Asn Asp Leu Arg Pro Glu
355 360 365

Ala Arg Pro Gly Gly Trp Val Gly Glu Trp Leu Ala Gln Tyr Pro Asp
370 375 380

Asn Tyr Glu Val Val Asn Thr Leu Ala Arg Gln Ile Lys Asp Ile Trp
385 390 395 400

Lys Asn Asn Gln His His Lys Asp Gly Gly Glu Pro Tyr Lys Leu Ala
405 410 415

Gln Arg Leu Ala Met Leu Ala His Glu Ile Asp Ala Val Pro Ala Trp
420 425 430

Asn Cys Lys Ser Gly Lys Asp Arg Thr Gly Met Met Asp Ser Glu Ile
435 440 445

Lys Gly Glu Ile Ile Ser Leu His Gln Thr His Met Leu Ser Ala Pro
450 455 460

Gly Ser Leu Pro Asp Ser Gly Gly Gln Lys Ile Phe Gln Lys Val Leu
465 470 475 480

Leu Asn Ser Gly Asn Leu Glu Ile Gln Lys Gln Asn Thr Gly Gly Ala
485 490 495

Gly Asn Lys Val Met Lys Asn Leu Ser Pro Glu Val Leu Asn Leu Ser
500 505 510

Tyr Gln Lys Arg Val Gly Asp Glu Asn Ile Trp Gln Ser Val Lys Gly
515 520 525

Ile Ser Ser Leu Ile Thr Ser Arg Ser Cys Gly Leu Val Pro Arg Gly
530 535 540

Ser Gly Pro Gly Ser Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu
545 550 555 560

Asp Trp Asp Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu
565 570 575

Lys Glu His Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys
580 585 590

Thr Val Ser Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln
595 600 605

Thr Ala Leu Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly
610 615 620

Thr Asn Pro Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn
625 630 635 640

Val Ala Gln Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr
645 650 655

Thr Ala Ala Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile
660 665 670

Ala Asp Gly Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser
675 680 685

Ile Ala Leu Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly
690 695 700

Glu Leu Val Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile
705 710 715 720

Ile Asn Leu Phe Gln Val Val His Asn Ser Tyr Asn Arg Ser Ala Tyr
725 730 735

Ser Pro Gly His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val
740 745 750

Ser Trp Asn Thr Val Arg Ser Thr Met Ser Tyr Thr Asn Asp Lys Ile
755 760 765

Leu Ile Leu Tyr Phe Asn Lys Leu Tyr Lys Lys Ile Lys Asp Asn Ser
770 775 780

Ile Leu Asp Met Arg Tyr Glu Asn Asn Lys Phe Ile Asp Ile Ser Gly
785 790 795 800

Tyr Gly Ser Asn Ile Ser Ile Asn Gly Asp Val Tyr Ile Tyr Ser Thr
805 810 815

Asn Arg Asn Gln Phe Gly Ile Tyr Ser Ser Lys Pro Ser Glu Val Asn
820 825 830

Ile Ala Gln Asn Asn Asp Ile Ile Tyr Asn Gly Arg Tyr Gln Asn Phe
835 840 845

Ser Ile Ser Phe Trp Val Arg Ile Pro Lys Tyr Phe Asn Lys Val Asn
850 855 860

Leu Asn Asn Glu Tyr Thr Ile Ile Asp Cys Ile Arg Asn Asn Asn Ser
865 870 875 880

Gly Trp Lys Ile Ser Leu Asn Tyr Asn Lys Ile Ile Trp Thr Leu Gln
885 890 895

Asp Thr Ala Gly Asn Asn Gln Lys Leu Val Phe Asn Tyr Thr Gln Met
900 905 910

Ile Ser Ile Ser Asp Tyr Ile Asn Lys Trp Ile Phe Val Thr Ile Thr
915 920 925

Asn Asn Arg Leu Gly Asn Ser Arg Ile Tyr Ile Asn Gly Asn Leu Ile
930 935 940

Asp Glu Lys Ser Ile Ser Asn Leu Gly Asp Ile His Val Ser Asp Asn
945 950 955 960

Ile Leu Phe Lys Ile Val Gly Cys Asn Asp Thr Arg Tyr Val Gly Ile
965 970 975

Arg Tyr Phe Lys Val Phe Asp Thr Glu Leu Gly Lys Thr Glu Ile Glu
980 985 990

Thr Leu Tyr Ser Asp Glu Pro Asp Pro Ser Ile Leu Lys Asp Phe Trp
995 1000 1005

Gly Asn Tyr Leu Leu Tyr Asn Lys Arg Tyr Tyr Leu Leu Asn Leu

1010 1015 1020

Leu Arg Thr Asp Lys Ser Ile Thr Gln Asn Ser Asn Phe Leu Asn
1025 1030 1035

Ile Asn Gln Gln Arg Gly Val Tyr Gln Lys Pro Asn Ile Phe Ser
1040 1045 1050

Asn Thr Arg Leu Tyr Thr Gly Val Glu Val Ile Ile Arg Lys Asn
1055 1060 1065

Gly Ser Thr Asp Ile Ser Asn Thr Asp Asn Phe Val Arg Lys Asn
1070 1075 1080

Asp Leu Ala Tyr Ile Asn Val Val Asp Arg Asp Val Glu Tyr Arg
1085 1090 1095

Leu Tyr Ala Asp Ile Ser Ile Ala Lys Pro Glu Lys Ile Ile Lys
1100 1105 1110

Leu Ile Arg Thr Ser Asn Ser Asn Asn Ser Leu Gly Gln Ile Ile
1115 1120 1125

Val Met Asp Ser Ile Gly Asn Asn Cys Thr Met Asn Phe Gln Asn
1130 1135 1140

Asn Asn Gly Gly Asn Ile Gly Leu Leu Gly Phe His Ser Asn Asn
1145 1150 1155

Leu Val Ala Ser Ser Trp Tyr Tyr Asn Asn Ile Arg Lys Asn Thr
1160 1165 1170

Ser Ser Asn Gly Cys Phe Trp Ser Phe Ile Ser Lys Glu His Gly
1175 1180 1185

Trp Gln Glu Asn
1190

<210> 24

<211> 1192

<212> PRT

<213> Protein sequence for SigD, factor Xa linker, diphtheria toxin translocation domain, with BoNT/F-HC

<400> 24

Met Gln Ile Leu Ser Gly Gln Gly Lys Ala Pro Ala Lys Ala Pro Asp
1 5 10 15

Ala Arg Pro Glu Ile Ile Val Leu Arg Glu Pro Gly Ala Thr Trp Gly
20 25 30

Asn Tyr Leu Gln His Gln Lys Ala Ser Asn His Ser Leu His Asn Leu
35 40 45

Tyr Asn Leu Gln Arg Asp Leu Leu Thr Val Ala Ala Thr Val Leu Gly
50 55 60

Lys Gln Asp Pro Val Leu Thr Ser Met Ala Asn Gln Met Glu Leu Ala
65 70 75 80

Lys Val Lys Ala Asp Arg Pro Ala Thr Lys Gln Glu Glu Ala Ala Ala
85 90 95

Lys Ala Leu Lys Lys Asn Leu Ile Glu Leu Ile Ala Ala Arg Thr Gln
100 105 110

Gln Gln Asp Gly Leu Pro Ala Lys Glu Ala His Arg Phe Ala Ala Val
115 120 125

Ala Phe Arg Asp Ala Gln Val Lys Gln Leu Asn Asn Gln Pro Trp Gln
130 135 140

Thr Ile Lys Asn Thr Leu Thr His Asn Gly His His Tyr Thr Asn Thr
145 150 155 160

Gln Leu Pro Ala Ala Glu Met Lys Ile Gly Ala Lys Asp Ile Phe Pro
165 170 175

Ser Ala Tyr Glu Gly Lys Gly Val Cys Ser Trp Asp Thr Lys Asn Ile
180 185 190

His His Ala Asn Asn Leu Trp Met Ser Thr Val Ser Val His Glu Asp
195 200 205

Gly Lys Asp Lys Thr Leu Phe Phe Asp Gly Ile Arg His Gly Val Leu
210 215 220

Ser Pro Tyr His Glu Lys Asp Pro Leu Leu Arg His Val Gly Ala Glu
225 230 235 240

Asn Lys Ala Lys Glu Val Leu Thr Ala Ala Leu Phe Ser Lys Pro Glu
245 250 255

Leu Leu Asn Lys Ala Leu Ala Gly Glu Ala Val Ser Leu Lys Leu Val
260 265 270

Ser Val Gly Leu Leu Thr Ala Ser Asn Ile Phe Gly Lys Glu Gly Thr

275 280 285

Met Val Glu Asp Gln Met Arg Ala Trp Gln Ser Leu Thr Gln Pro Gly
290 295 300

Lys Met Ile His Leu Lys Ile Arg Asn Lys Asp Gly Asp Leu Gln Thr
305 310 315 320

Val Lys Ile Lys Pro Asp Val Val Ala Ala Phe Asn Val Gly Val Asn
325 330 335

Glu Leu Ala Leu Lys Leu Gly Phe Gly Leu Lys Ala Ser Asp Ser Tyr
340 345 350

Asn Ala Glu Ala Leu His Gln Leu Leu Gly Asn Asp Leu Arg Pro Glu
355 360 365

Ala Arg Pro Gly Gly Trp Val Gly Glu Trp Leu Ala Gln Tyr Pro Asp
370 375 380

Asn Tyr Glu Val Val Asn Thr Leu Ala Arg Gln Ile Lys Asp Ile Trp
385 390 395 400

Lys Asn Asn Gln His His Lys Asp Gly Gly Glu Pro Tyr Lys Leu Ala
405 410 415

Gln Arg Leu Ala Met Leu Ala His Glu Ile Asp Ala Val Pro Ala Trp
420 425 430

Asn Cys Lys Ser Gly Lys Asp Arg Thr Gly Met Met Asp Ser Glu Ile
435 440 445

Lys Gly Glu Ile Ile Ser Leu His Gln Thr His Met Leu Ser Ala Pro
450 455 460

Gly Ser Leu Pro Asp Ser Gly Gly Gln Lys Ile Phe Gln Lys Val Leu
465 470 475 480

Leu Asn Ser Gly Asn Leu Glu Ile Gln Lys Gln Asn Thr Gly Gly Ala
485 490 495

Gly Asn Lys Val Met Lys Asn Leu Ser Pro Glu Val Leu Asn Leu Ser
500 505 510

Tyr Gln Lys Arg Val Gly Asp Glu Asn Ile Trp Gln Ser Val Lys Gly
515 520 525

Ile Ser Ser Leu Ile Thr Ser Arg Ser Cys Gly Ile Glu Gly Arg Ala
530 535 540

Pro Gly Pro Gly Ser Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu
545 550 555 560

Asp Trp Asp Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu
565 570 575

Lys Glu His Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys
580 585 590

Thr Val Ser Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln
595 600 605

Thr Ala Leu Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly
610 615 620

Thr Asn Pro Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn
625 630 635 640

Val Ala Gln Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr
645 650 655

Thr Ala Ala Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile
660 665 670

Ala Asp Gly Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser
675 680 685

Ile Ala Leu Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly
690 695 700

Glu Leu Val Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile
705 710 715 720

Ile Asn Leu Phe Gln Val Val His Asn Ser Tyr Asn Arg Ser Ala Tyr
725 730 735

Ser Pro Gly His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val
740 745 750

Ser Trp Asn Thr Val Arg Ser Thr Met Ser Tyr Thr Asn Asp Lys Ile
755 760 765

Leu Ile Leu Tyr Phe Asn Lys Leu Tyr Lys Lys Ile Lys Asp Asn Ser
770 775 780

Ile Leu Asp Met Arg Tyr Glu Asn Asn Lys Phe Ile Asp Ile Ser Gly
785 790 795 800

Tyr Gly Ser Asn Ile Ser Ile Asn Gly Asp Val Tyr Ile Tyr Ser Thr
805 810 815

Asn Arg Asn Gln Phe Gly Ile Tyr Ser Ser Lys Pro Ser Glu Val Asn
820 825 830

Ile Ala Gln Asn Asn Asp Ile Ile Tyr Asn Gly Arg Tyr Gln Asn Phe
835 840 845

Ser Ile Ser Phe Trp Val Arg Ile Pro Lys Tyr Phe Asn Lys Val Asn
850 855 860

Leu Asn Asn Glu Tyr Thr Ile Ile Asp Cys Ile Arg Asn Asn Asn Ser
865 870 875 880

Gly Trp Lys Ile Ser Leu Asn Tyr Asn Lys Ile Ile Trp Thr Leu Gln
885 890 895

Asp Thr Ala Gly Asn Asn Gln Lys Leu Val Phe Asn Tyr Thr Gln Met
900 905 910

Ile Ser Ile Ser Asp Tyr Ile Asn Lys Trp Ile Phe Val Thr Ile Thr
915 920 925

Asn Asn Arg Leu Gly Asn Ser Arg Ile Tyr Ile Asn Gly Asn Leu Ile
930 935 940

Asp Glu Lys Ser Ile Ser Asn Leu Gly Asp Ile His Val Ser Asp Asn
945 950 955 960

Ile Leu Phe Lys Ile Val Gly Cys Asn Asp Thr Arg Tyr Val Gly Ile
965 970 975

Arg Tyr Phe Lys Val Phe Asp Thr Glu Leu Gly Lys Thr Glu Ile Glu
980 985 990

Thr Leu Tyr Ser Asp Glu Pro Asp Pro Ser Ile Leu Lys Asp Phe Trp
995 1000 1005

Gly Asn Tyr Leu Leu Tyr Asn Lys Arg Tyr Tyr Leu Leu Asn Leu
1010 1015 1020

Leu Arg Thr Asp Lys Ser Ile Thr Gln Asn Ser Asn Phe Leu Asn
1025 1030 1035

Ile Asn Gln Gln Arg Gly Val Tyr Gln Lys Pro Asn Ile Phe Ser
1040 1045 1050

Asn Thr Arg Leu Tyr Thr Gly Val Glu Val Ile Ile Arg Lys Asn
1055 1060 1065

Gly Ser Thr Asp Ile Ser Asn Thr Asp Asn Phe Val Arg Lys Asn

1070 1075 1080

Asp Leu Ala Tyr Ile Asn Val Val Asp Arg Asp Val Glu Tyr Arg
1085 1090 1095

Leu Tyr Ala Asp Ile Ser Ile Ala Lys Pro Glu Lys Ile Ile Lys
1100 1105 1110

Leu Ile Arg Thr Ser Asn Ser Asn Asn Ser Leu Gly Gln Ile Ile
1115 1120 1125

Val Met Asp Ser Ile Gly Asn Asn Cys Thr Met Asn Phe Gln Asn
1130 1135 1140

Asn Asn Gly Gly Asn Ile Gly Leu Leu Gly Phe His Ser Asn Asn
1145 1150 1155

Leu Val Ala Ser Ser Trp Tyr Tyr Asn Asn Ile Arg Lys Asn Thr
1160 1165 1170

Ser Ser Asn Gly Cys Phe Trp Ser Phe Ile Ser Lys Glu His Gly
1175 1180 1185

Trp Gln Glu Asn
1190

<210> 25

<211> 999

<212> PRT

<213> Protein sequence for YopT, factor Xa linker, diphtheria translocation domain, TeNT- HC

<400> 25

Met Asn Ser Ile His Gly His Tyr His Ile Gln Leu Ser Asn Tyr Ser
1 5 10 15

Ala Gly Glu Asn Leu Gln Ser Ala Thr Leu Thr Glu Gly Val Ile Gly
20 25 30

Ala His Arg Val Lys Val Glu Thr Ala Leu Ser His Ser Asn Leu Gln
35 40 45

Lys Lys Leu Ser Ala Thr Ile Lys His Asn Gln Ser Gly Arg Ser Met
50 55 60

Leu Asp Arg Lys Leu Thr Ser Asp Gly Lys Ala Asn Gln Arg Ser Ser
65 70 75 80

```
Phe Thr Phe Ser Met Ile Met Tyr Arg Met Ile His Phe Val Leu Ser
                85                  90                  95

Thr Arg Val Pro Ala Val Arg Glu Ser Val Ala Asn Tyr Gly Gly Asn
            100                 105                 110

Ile Asn Phe Lys Phe Ala Gln Thr Lys Gly Ala Phe Leu His Lys Ile
        115                 120                 125

Ile Lys His Ser Asp Thr Ala Ser Gly Val Cys Glu Ala Leu Cys Ala
    130                 135                 140

His Trp Ile Arg Ser His Ala Gln Gly Gln Ser Leu Phe Asp Gln Leu
145                 150                 155                 160

Tyr Val Gly Gly Arg Lys Gly Lys Phe Gln Ile Asp Thr Leu Tyr Ser
                165                 170                 175

Ile Lys Gln Leu Gln Ile Asp Gly Cys Lys Ala Asp Val Asp Gln Asp
            180                 185                 190

Glu Val Thr Leu Asp Trp Phe Lys Lys Asn Gly Ile Ser Glu Arg Met
        195                 200                 205

Ile Glu Arg His Cys Leu Leu Arg Pro Val Asp Val Thr Gly Thr Thr
    210                 215                 220

Glu Ser Glu Gly Leu Asp Gln Leu Leu Asn Ala Ile Leu Asp Thr His
225                 230                 235                 240

Gly Ile Gly Tyr Gly Tyr Lys Lys Ile His Leu Ser Gly Gln Met Ser
                245                 250                 255

Ala His Ala Ile Ala Ala Tyr Val Asn Glu Lys Ser Gly Val Thr Phe
            260                 265                 270

Phe Asp Pro Asn Phe Gly Glu Phe His Phe Ser Asp Lys Glu Lys Phe
        275                 280                 285

Arg Lys Trp Phe Thr Asn Ser Phe Trp Gly Asn Ser Met Tyr His Tyr
    290                 295                 300

Pro Leu Gly Val Gly Gln Arg Phe Arg Val Leu Thr Phe Asp Ser Lys
305                 310                 315                 320

Glu Val Arg Ser Cys Gly Ile Glu Gly Arg Ala Pro Gly Pro Gly Ser
                325                 330                 335

Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp Val Ile
```

```
            340                 345                 350

Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly Pro
        355                 360                 365

Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu Glu
    370                 375                 380

Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu His
385                 390                 395                 400

Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val Phe
                405                 410                 415

Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val Ile
            420                 425                 430

Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu Ser
        435                 440                 445

Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala Val
    450                 455                 460

His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu Ser Ser
465                 470                 475                 480

Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp Ile
                485                 490                 495

Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe Gln
            500                 505                 510

Val Val His Asn Ser Tyr Asn Arg Ser Ala Tyr Ser Pro Gly His Lys
        515                 520                 525

Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr Val
    530                 535                 540

Arg Ser Lys Asn Leu Asp Cys Trp Val Asp Asn Glu Glu Asp Ile Asp
545                 550                 555                 560

Val Ile Leu Lys Lys Ser Thr Ile Leu Asn Leu Asp Ile Asn Asn Asp
                565                 570                 575

Ile Ile Ser Asp Ile Ser Gly Phe Asn Ser Ser Val Ile Thr Tyr Pro
            580                 585                 590

Asp Ala Gln Leu Val Pro Gly Ile Asn Gly Lys Ala Ile His Leu Val
        595                 600                 605
```

Asn Asn Glu Ser Ser Glu Val Ile Val His Lys Ala Met Asp Ile Glu
610 615 620

Tyr Asn Asp Met Phe Asn Asn Phe Thr Val Ser Phe Trp Leu Arg Val
625 630 635 640

Pro Lys Val Ser Ala Ser His Leu Glu Gln Tyr Gly Thr Asn Glu Tyr
645 650 655

Ser Ile Ile Ser Ser Met Lys Lys His Ser Leu Ser Ile Gly Ser Gly
660 665 670

Trp Ser Val Ser Leu Lys Gly Asn Asn Leu Ile Trp Thr Leu Lys Asp
675 680 685

Ser Ala Gly Glu Val Arg Gln Ile Thr Phe Arg Asp Leu Pro Asp Lys
690 695 700

Phe Asn Ala Tyr Leu Ala Asn Lys Trp Val Phe Ile Thr Ile Thr Asn
705 710 715 720

Asp Arg Leu Ser Ser Ala Asn Leu Tyr Ile Asn Gly Val Leu Met Gly
725 730 735

Ser Ala Glu Ile Thr Gly Leu Gly Ala Ile Arg Glu Asp Asn Asn Ile
740 745 750

Thr Leu Lys Leu Asp Arg Cys Asn Asn Asn Asn Gln Tyr Val Ser Ile
755 760 765

Asp Lys Phe Arg Ile Phe Cys Lys Ala Leu Asn Pro Lys Glu Ile Glu
770 775 780

Lys Leu Tyr Thr Ser Tyr Leu Ser Ile Thr Phe Leu Arg Asp Phe Trp
785 790 795 800

Gly Asn Pro Leu Arg Tyr Asp Thr Glu Tyr Tyr Leu Ile Pro Val Ala
805 810 815

Ser Ser Ser Lys Asp Val Gln Leu Lys Asn Ile Thr Asp Tyr Met Tyr
820 825 830

Leu Thr Asn Ala Pro Ser Tyr Thr Asn Gly Lys Leu Asn Ile Tyr Tyr
835 840 845

Arg Arg Leu Tyr Asn Gly Leu Lys Phe Ile Ile Lys Arg Tyr Thr Pro
850 855 860

Asn Asn Glu Ile Asp Ser Phe Val Lys Ser Gly Asp Phe Ile Lys Leu
865 870 875 880

Tyr Val Ser Tyr Asn Asn Asn Glu His Ile Val Gly Tyr Pro Lys Asp
885 890 895

Gly Asn Ala Phe Asn Asn Leu Asp Arg Ile Leu Arg Val Gly Tyr Asn
900 905 910

Ala Pro Gly Ile Pro Leu Tyr Lys Lys Met Glu Ala Val Lys Leu Arg
915 920 925

Asp Leu Lys Thr Tyr Ser Val Gln Leu Lys Leu Tyr Asp Asp Lys Asn
930 935 940

Ala Ser Leu Gly Leu Val Gly Thr His Asn Gly Gln Ile Gly Asn Asp
945 950 955 960

Pro Asn Arg Asp Ile Leu Ile Ala Ser Asn Trp Tyr Phe Asn His Leu
965 970 975

Lys Asp Lys Ile Leu Gly Cys Asp Trp Tyr Phe Val Pro Thr Asp Glu
980 985 990

Gly Trp Thr Asn Asp Leu Gln
995

<210> 26

<211> 979

<212> PRT

<213> Protein sequence for YopT, factor Xa linker, diphtheria toxin translocation domain, with BoNT/F-HC

<400> 26

Met Asn Ser Ile His Gly His Tyr His Ile Gln Leu Ser Asn Tyr Ser
1 5 10 15

Ala Gly Glu Asn Leu Gln Ser Ala Thr Leu Thr Glu Gly Val Ile Gly
20 25 30

Ala His Arg Val Lys Val Glu Thr Ala Leu Ser His Ser Asn Leu Gln
35 40 45

Lys Lys Leu Ser Ala Thr Ile Lys His Asn Gln Ser Gly Arg Ser Met
50 55 60

Leu Asp Arg Lys Leu Thr Ser Asp Gly Lys Ala Asn Gln Arg Ser Ser
65 70 75 80

Phe Thr Phe Ser Met Ile Met Tyr Arg Met Ile His Phe Val Leu Ser
85 90 95

Thr Arg Val Pro Ala Val Arg Glu Ser Val Ala Asn Tyr Gly Gly Asn
100 105 110

Ile Asn Phe Lys Phe Ala Gln Thr Lys Gly Ala Phe Leu His Lys Ile
115 120 125

Ile Lys His Ser Asp Thr Ala Ser Gly Val Cys Glu Ala Leu Cys Ala
130 135 140

His Trp Ile Arg Ser His Ala Gln Gly Gln Ser Leu Phe Asp Gln Leu
145 150 155 160

Tyr Val Gly Gly Arg Lys Gly Lys Phe Gln Ile Asp Thr Leu Tyr Ser
165 170 175

Ile Lys Gln Leu Gln Ile Asp Gly Cys Lys Ala Asp Val Asp Gln Asp
180 185 190

Glu Val Thr Leu Asp Trp Phe Lys Lys Asn Gly Ile Ser Glu Arg Met
195 200 205

Ile Glu Arg His Cys Leu Leu Arg Pro Val Asp Val Thr Gly Thr Thr
210 215 220

Glu Ser Glu Gly Leu Asp Gln Leu Leu Asn Ala Ile Leu Asp Thr His
225 230 235 240

Gly Ile Gly Tyr Gly Tyr Lys Lys Ile His Leu Ser Gly Gln Met Ser
245 250 255

Ala His Ala Ile Ala Ala Tyr Val Asn Glu Lys Ser Gly Val Thr Phe
260 265 270

Phe Asp Pro Asn Phe Gly Glu Phe His Phe Ser Asp Lys Glu Lys Phe
275 280 285

Arg Lys Trp Phe Thr Asn Ser Phe Trp Gly Asn Ser Met Tyr His Tyr
290 295 300

Pro Leu Gly Val Gly Gln Arg Phe Arg Val Leu Thr Phe Asp Ser Lys
305 310 315 320

Glu Val Arg Ser Cys Gly Ile Glu Gly Arg Ala Pro Gly Pro Gly Ser
325 330 335

Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp Val Ile
340 345 350

Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly Pro
355 360 365

Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu Glu
370 375 380

Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu His
385 390 395 400

Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val Phe
405 410 415

Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val Ile
420 425 430

Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu Ser
435 440 445

Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala Val
450 455 460

His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu Ser Ser
465 470 475 480

Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp Ile
485 490 495

Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe Gln
500 505 510

Val Val His Asn Ser Tyr Asn Arg Ser Ala Tyr Ser Pro Gly His Lys
515 520 525

Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr Val
530 535 540

Arg Ser Thr Met Ser Tyr Thr Asn Asp Lys Ile Leu Ile Leu Tyr Phe
545 550 555 560

Asn Lys Leu Tyr Lys Lys Ile Lys Asp Asn Ser Ile Leu Asp Met Arg
565 570 575

Tyr Glu Asn Asn Lys Phe Ile Asp Ile Ser Gly Tyr Gly Ser Asn Ile
580 585 590

Ser Ile Asn Gly Asp Val Tyr Ile Tyr Ser Thr Asn Arg Asn Gln Phe
595 600 605

Gly Ile Tyr Ser Ser Lys Pro Ser Glu Val Asn Ile Ala Gln Asn Asn

610 615 620

Asp Ile Ile Tyr Asn Gly Arg Tyr Gln Asn Phe Ser Ile Ser Phe Trp
625 630 635 640

Val Arg Ile Pro Lys Tyr Phe Asn Lys Val Asn Leu Asn Asn Glu Tyr
645 650 655

Thr Ile Ile Asp Cys Ile Arg Asn Asn Asn Ser Gly Trp Lys Ile Ser
660 665 670

Leu Asn Tyr Asn Lys Ile Ile Trp Thr Leu Gln Asp Thr Ala Gly Asn
675 680 685

Asn Gln Lys Leu Val Phe Asn Tyr Thr Gln Met Ile Ser Ile Ser Asp
690 695 700

Tyr Ile Asn Lys Trp Ile Phe Val Thr Ile Thr Asn Asn Arg Leu Gly
705 710 715 720

Asn Ser Arg Ile Tyr Ile Asn Gly Asn Leu Ile Asp Glu Lys Ser Ile
725 730 735

Ser Asn Leu Gly Asp Ile His Val Ser Asp Asn Ile Leu Phe Lys Ile
740 745 750

Val Gly Cys Asn Asp Thr Arg Tyr Val Gly Ile Arg Tyr Phe Lys Val
755 760 765

Phe Asp Thr Glu Leu Gly Lys Thr Glu Ile Glu Thr Leu Tyr Ser Asp
770 775 780

Glu Pro Asp Pro Ser Ile Leu Lys Asp Phe Trp Gly Asn Tyr Leu Leu
785 790 795 800

Tyr Asn Lys Arg Tyr Tyr Leu Leu Asn Leu Leu Arg Thr Asp Lys Ser
805 810 815

Ile Thr Gln Asn Ser Asn Phe Leu Asn Ile Asn Gln Gln Arg Gly Val
820 825 830

Tyr Gln Lys Pro Asn Ile Phe Ser Asn Thr Arg Leu Tyr Thr Gly Val
835 840 845

Glu Val Ile Ile Arg Lys Asn Gly Ser Thr Asp Ile Ser Asn Thr Asp
850 855 860

Asn Phe Val Arg Lys Asn Asp Leu Ala Tyr Ile Asn Val Val Asp Arg
865 870 875 880

Asp Val Glu Tyr Arg Leu Tyr Ala Asp Ile Ser Ile Ala Lys Pro Glu
885 890 895

Lys Ile Ile Lys Leu Ile Arg Thr Ser Asn Ser Asn Asn Ser Leu Gly
900 905 910

Gln Ile Ile Val Met Asp Ser Ile Gly Asn Asn Cys Thr Met Asn Phe
915 920 925

Gln Asn Asn Asn Gly Gly Asn Ile Gly Leu Leu Gly Phe His Ser Asn
930 935 940

Asn Leu Val Ala Ser Ser Trp Tyr Tyr Asn Asn Ile Arg Lys Asn Thr
945 950 955 960

Ser Ser Asn Gly Cys Phe Trp Ser Phe Ile Ser Lys Glu His Gly Trp
965 970 975

Gln Glu Asn

<210> 27

<211> 810

<212> PRT

<213> Protein sequence for SpiC, thrombin linker, diphtheria translocation domain, TeNT-HC

<400> 27

Met Ser Glu Glu Gly Phe Met Leu Ala Val Leu Lys Gly Ile Pro Leu
1 5 10 15

Ile Gln Asp Ile Arg Ala Glu Gly Asn Ser Arg Ser Trp Ile Met Thr
20 25 30

Ile Asp Gly His Pro Ala Arg Gly Glu Ile Phe Ser Glu Ala Phe Ser
35 40 45

Ile Ser Leu Phe Leu Asn Asp Leu Glu Ser Leu Pro Lys Pro Cys Leu
50 55 60

Ala Tyr Val Thr Leu Leu Leu Ala Ala His Pro Asp Val His Asp Tyr
65 70 75 80

Ala Ile Gln Leu Thr Ala Asp Gly Gly Trp Leu Asn Gly Tyr Tyr Thr
85 90 95

Thr Ser Ser Ser Ser Glu Leu Ile Ala Ile Glu Ile Glu Lys His Leu

100 105 110

Ala Leu Thr Cys Ile Leu Lys Asn Val Ile Arg Asn His His Lys Leu
115 120 125

Tyr Ser Gly Gly Val Arg Ser Cys Gly Leu Val Pro Arg Gly Ser Gly
130 135 140

Pro Gly Ser Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp
145 150 155 160

Asp Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu
165 170 175

His Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val
180 185 190

Ser Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala
195 200 205

Leu Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn
210 215 220

Pro Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala
225 230 235 240

Gln Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala
245 250 255

Ala Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp
260 265 270

Gly Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala
275 280 285

Leu Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu
290 295 300

Val Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn
305 310 315 320

Leu Phe Gln Val Val His Asn Ser Tyr Asn Arg Ser Ala Tyr Ser Pro
325 330 335

Gly His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp
340 345 350

Asn Thr Val Arg Ser Lys Asn Leu Asp Cys Trp Val Asp Asn Glu Glu
355 360 365

Asp Ile Asp Val Ile Leu Lys Lys Ser Thr Ile Leu Asn Leu Asp Ile
370 375 380

Asn Asn Asp Ile Ile Ser Asp Ile Ser Gly Phe Asn Ser Ser Val Ile
385 390 395 400

Thr Tyr Pro Asp Ala Gln Leu Val Pro Gly Ile Asn Gly Lys Ala Ile
405 410 415

His Leu Val Asn Asn Glu Ser Ser Glu Val Ile Val His Lys Ala Met
420 425 430

Asp Ile Glu Tyr Asn Asp Met Phe Asn Asn Phe Thr Val Ser Phe Trp
435 440 445

Leu Arg Val Pro Lys Val Ser Ala Ser His Leu Glu Gln Tyr Gly Thr
450 455 460

Asn Glu Tyr Ser Ile Ile Ser Ser Met Lys Lys His Ser Leu Ser Ile
465 470 475 480

Gly Ser Gly Trp Ser Val Ser Leu Lys Gly Asn Asn Leu Ile Trp Thr
485 490 495

Leu Lys Asp Ser Ala Gly Glu Val Arg Gln Ile Thr Phe Arg Asp Leu
500 505 510

Pro Asp Lys Phe Asn Ala Tyr Leu Ala Asn Lys Trp Val Phe Ile Thr
515 520 525

Ile Thr Asn Asp Arg Leu Ser Ser Ala Asn Leu Tyr Ile Asn Gly Val
530 535 540

Leu Met Gly Ser Ala Glu Ile Thr Gly Leu Gly Ala Ile Arg Glu Asp
545 550 555 560

Asn Asn Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn Asn Asn Gln Tyr
565 570 575

Val Ser Ile Asp Lys Phe Arg Ile Phe Cys Lys Ala Leu Asn Pro Lys
580 585 590

Glu Ile Glu Lys Leu Tyr Thr Ser Tyr Leu Ser Ile Thr Phe Leu Arg
595 600 605

Asp Phe Trp Gly Asn Pro Leu Arg Tyr Asp Thr Glu Tyr Tyr Leu Ile
610 615 620

Pro Val Ala Ser Ser Ser Lys Asp Val Gln Leu Lys Asn Ile Thr Asp
625 630 635 640

Tyr Met Tyr Leu Thr Asn Ala Pro Ser Tyr Thr Asn Gly Lys Leu Asn
645 650 655

Ile Tyr Tyr Arg Arg Leu Tyr Asn Gly Leu Lys Phe Ile Ile Lys Arg
660 665 670

Tyr Thr Pro Asn Asn Glu Ile Asp Ser Phe Val Lys Ser Gly Asp Phe
675 680 685

Ile Lys Leu Tyr Val Ser Tyr Asn Asn Asn Glu His Ile Val Gly Tyr
690 695 700

Pro Lys Asp Gly Asn Ala Phe Asn Asn Leu Asp Arg Ile Leu Arg Val
705 710 715 720

Gly Tyr Asn Ala Pro Gly Ile Pro Leu Tyr Lys Lys Met Glu Ala Val
725 730 735

Lys Leu Arg Asp Leu Lys Thr Tyr Ser Val Gln Leu Lys Leu Tyr Asp
740 745 750

Asp Lys Asn Ala Ser Leu Gly Leu Val Gly Thr His Asn Gly Gln Ile
755 760 765

Gly Asn Asp Pro Asn Arg Asp Ile Leu Ile Ala Ser Asn Trp Tyr Phe
770 775 780

Asn His Leu Lys Asp Lys Ile Leu Gly Cys Asp Trp Tyr Phe Val Pro
785 790 795 800

Thr Asp Glu Gly Trp Thr Asn Asp Leu Gln
805 810

<210> 28

<211> 810

<212> PRT

<213> Protein sequence for SpiC, factor Xa linker, diphtheria translocation domain, TeNT- HC

<400> 28

Met Ser Glu Glu Gly Phe Met Leu Ala Val Leu Lys Gly Ile Pro Leu
1 5 10 15

Ile Gln Asp Ile Arg Ala Glu Gly Asn Ser Arg Ser Trp Ile Met Thr
20 25 30

```
Ile Asp Gly His Pro Ala Arg Gly Glu Ile Phe Ser Glu Ala Phe Ser
        35                  40                  45

Ile Ser Leu Phe Leu Asn Asp Leu Glu Ser Leu Pro Lys Pro Cys Leu
    50                  55                  60

Ala Tyr Val Thr Leu Leu Leu Ala Ala His Pro Asp Val His Asp Tyr
65                  70                  75                  80

Ala Ile Gln Leu Thr Ala Asp Gly Gly Trp Leu Asn Gly Tyr Tyr Thr
                85                  90                  95

Thr Ser Ser Ser Ser Glu Leu Ile Ala Ile Glu Ile Glu Lys His Leu
            100                 105                 110

Ala Leu Thr Cys Ile Leu Lys Asn Val Ile Arg Asn His His Lys Leu
        115                 120                 125

Tyr Ser Gly Gly Val Arg Ser Cys Gly Ile Glu Gly Arg Ala Pro Gly
    130                 135                 140

Pro Gly Ser Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp
145                 150                 155                 160

Asp Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu
                165                 170                 175

His Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val
            180                 185                 190

Ser Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala
        195                 200                 205

Leu Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn
    210                 215                 220

Pro Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala
225                 230                 235                 240

Gln Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala
                245                 250                 255

Ala Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp
            260                 265                 270

Gly Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala
        275                 280                 285

Leu Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu
    290                 295                 300
```

Val Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn
305 310 315 320

Leu Phe Gln Val Val His Asn Ser Tyr Asn Arg Ser Ala Tyr Ser Pro
325 330 335

Gly His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp
340 345 350

Asn Thr Val Arg Ser Lys Asn Leu Asp Cys Trp Val Asp Asn Glu Glu
355 360 365

Asp Ile Asp Val Ile Leu Lys Lys Ser Thr Ile Leu Asn Leu Asp Ile
370 375 380

Asn Asn Asp Ile Ile Ser Asp Ile Ser Gly Phe Asn Ser Ser Val Ile
385 390 395 400

Thr Tyr Pro Asp Ala Gln Leu Val Pro Gly Ile Asn Gly Lys Ala Ile
405 410 415

His Leu Val Asn Asn Glu Ser Ser Glu Val Ile Val His Lys Ala Met
420 425 430

Asp Ile Glu Tyr Asn Asp Met Phe Asn Asn Phe Thr Val Ser Phe Trp
435 440 445

Leu Arg Val Pro Lys Val Ser Ala Ser His Leu Glu Gln Tyr Gly Thr
450 455 460

Asn Glu Tyr Ser Ile Ile Ser Ser Met Lys Lys His Ser Leu Ser Ile
465 470 475 480

Gly Ser Gly Trp Ser Val Ser Leu Lys Gly Asn Asn Leu Ile Trp Thr
485 490 495

Leu Lys Asp Ser Ala Gly Glu Val Arg Gln Ile Thr Phe Arg Asp Leu
500 505 510

Pro Asp Lys Phe Asn Ala Tyr Leu Ala Asn Lys Trp Val Phe Ile Thr
515 520 525

Ile Thr Asn Asp Arg Leu Ser Ser Ala Asn Leu Tyr Ile Asn Gly Val
530 535 540

Leu Met Gly Ser Ala Glu Ile Thr Gly Leu Gly Ala Ile Arg Glu Asp
545 550 555 560

Asn Asn Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn Asn Asn Gln Tyr

565 570 575

Val Ser Ile Asp Lys Phe Arg Ile Phe Cys Lys Ala Leu Asn Pro Lys
580 585 590

Glu Ile Glu Lys Leu Tyr Thr Ser Tyr Leu Ser Ile Thr Phe Leu Arg
595 600 605

Asp Phe Trp Gly Asn Pro Leu Arg Tyr Asp Thr Glu Tyr Tyr Leu Ile
610 615 620

Pro Val Ala Ser Ser Ser Lys Asp Val Gln Leu Lys Asn Ile Thr Asp
625 630 635 640

Tyr Met Tyr Leu Thr Asn Ala Pro Ser Tyr Thr Asn Gly Lys Leu Asn
645 650 655

Ile Tyr Tyr Arg Arg Leu Tyr Asn Gly Leu Lys Phe Ile Ile Lys Arg
660 665 670

Tyr Thr Pro Asn Asn Glu Ile Asp Ser Phe Val Lys Ser Gly Asp Phe
675 680 685

Ile Lys Leu Tyr Val Ser Tyr Asn Asn Asn Glu His Ile Val Gly Tyr
690 695 700

Pro Lys Asp Gly Asn Ala Phe Asn Asn Leu Asp Arg Ile Leu Arg Val
705 710 715 720

Gly Tyr Asn Ala Pro Gly Ile Pro Leu Tyr Lys Lys Met Glu Ala Val
725 730 735

Lys Leu Arg Asp Leu Lys Thr Tyr Ser Val Gln Leu Lys Leu Tyr Asp
740 745 750

Asp Lys Asn Ala Ser Leu Gly Leu Val Gly Thr His Asn Gly Gln Ile
755 760 765

Gly Asn Asp Pro Asn Arg Asp Ile Leu Ile Ala Ser Asn Trp Tyr Phe
770 775 780

Asn His Leu Lys Asp Lys Ile Leu Gly Cys Asp Trp Tyr Phe Val Pro
785 790 795 800

Thr Asp Glu Gly Trp Thr Asn Asp Leu Gln
805 810

<210> 29

<211> 393

<212> PRT

<213> Protein sequence for SpiC fused to a domain consisting the N-terminal 254 residues from Bacillus anthracis lethal factor capable of interacting with protective antigen

<400> 29

Met Ser Glu Glu Gly Phe Met Leu Ala Val Leu Lys Gly Ile Pro Leu
1 5 10 15

Ile Gln Asp Ile Arg Ala Glu Gly Asn Ser Arg Ser Trp Ile Met Thr
20 25 30

Ile Asp Gly His Pro Ala Arg Gly Glu Ile Phe Ser Glu Ala Phe Ser
35 40 45

Ile Ser Leu Phe Leu Asn Asp Leu Glu Ser Leu Pro Lys Pro Cys Leu
50 55 60

Ala Tyr Val Thr Leu Leu Leu Ala Ala His Pro Asp Val His Asp Tyr
65 70 75 80

Ala Ile Gln Leu Thr Ala Asp Gly Gly Trp Leu Asn Gly Tyr Tyr Thr
85 90 95

Thr Ser Ser Ser Ser Glu Leu Ile Ala Ile Glu Ile Glu Lys His Leu
100 105 110

Ala Leu Thr Cys Ile Leu Lys Asn Val Ile Arg Asn His His Lys Leu
115 120 125

Tyr Ser Gly Gly Val Met Asn Ile Lys Lys Glu Phe Ile Lys Val Ile
130 135 140

Ser Met Ser Cys Leu Val Thr Ala Ile Thr Leu Ser Gly Pro Val Phe
145 150 155 160

Ile Pro Leu Val Gln Gly Ala Gly Gly His Gly Asp Val Gly Met His
165 170 175

Val Lys Glu Lys Glu Lys Asn Lys Asp Glu Asn Lys Arg Lys Asp Glu
180 185 190

Glu Arg Asn Lys Thr Gln Glu Glu His Leu Lys Glu Ile Met Lys His
195 200 205

Ile Val Lys Ile Glu Val Lys Gly Glu Glu Ala Val Lys Lys Glu Ala
210 215 220

Ala Glu Lys Leu Leu Glu Lys Val Pro Ser Asp Val Leu Glu Met Tyr
225 230 235 240

Lys Ala Ile Gly Gly Lys Ile Tyr Ile Val Asp Gly Asp Ile Thr Lys
245 250 255

His Ile Ser Leu Glu Ala Leu Ser Glu Asp Lys Lys Lys Ile Lys Asp
260 265 270

Ile Tyr Gly Lys Asp Ala Leu Leu His Glu His Tyr Val Tyr Ala Lys
275 280 285

Glu Gly Tyr Glu Pro Val Leu Val Ile Gln Ser Ser Glu Asp Tyr Val
290 295 300

Glu Asn Thr Glu Lys Ala Leu Asn Val Tyr Tyr Glu Ile Gly Lys Ile
305 310 315 320

Leu Ser Arg Asp Ile Leu Ser Lys Ile Asn Gln Pro Tyr Gln Lys Phe
325 330 335

Leu Asp Val Leu Asn Thr Ile Lys Asn Ala Ser Asp Ser Asp Gly Gln
340 345 350

Asp Leu Leu Phe Thr Asn Gln Leu Lys Glu His Pro Thr Asp Phe Ser
355 360 365

Val Glu Phe Leu Glu Gln Asn Ser Asn Glu Val Gln Glu Val Phe Ala
370 375 380

Lys Ala Phe Ala Tyr Tyr Ile Glu Pro
385 390

<210> 30

<211> 764

<212> PRT

<213> Protein sequence of Bacillus anthracis protective antigen

<400> 30

Met Lys Lys Arg Lys Val Leu Ile Pro Leu Met Ala Leu Ser Thr Ile
1 5 10 15

Leu Val Ser Ser Thr Gly Asn Leu Glu Val Ile Gln Ala Glu Val Lys
20 25 30

Gln Glu Asn Arg Leu Leu Asn Glu Ser Glu Ser Ser Ser Gln Gly Leu
35 40 45

```
Leu Gly Tyr Tyr Phe Ser Asp Leu Asn Phe Gln Ala Pro Met Val Val
    50                  55                  60

Thr Ser Ser Thr Thr Gly Asp Leu Ser Ile Pro Ser Ser Glu Leu Glu
65                  70                  75                  80

Asn Ile Pro Ser Glu Asn Gln Tyr Phe Gln Ser Ala Ile Trp Ser Gly
                85                  90                  95

Phe Ile Lys Val Lys Lys Ser Asp Glu Tyr Thr Phe Ala Thr Ser Ala
            100                 105                 110

Asp Asn His Val Thr Met Trp Val Asp Asp Gln Glu Val Ile Asn Lys
        115                 120                 125

Ala Ser Asn Ser Asn Lys Ile Arg Leu Glu Lys Gly Arg Leu Tyr Gln
    130                 135                 140

Ile Lys Ile Gln Tyr Gln Arg Glu Asn Pro Thr Glu Lys Gly Leu Asp
145                 150                 155                 160

Phe Lys Leu Tyr Trp Thr Asp Ser Gln Asn Lys Lys Glu Val Ile Ser
                165                 170                 175

Ser Asp Asn Leu Gln Leu Pro Glu Leu Lys Gln Lys Ser Ser Asn Ser
            180                 185                 190

Arg Lys Lys Arg Ser Thr Ser Ala Gly Pro Thr Val Pro Asp Arg Asp
        195                 200                 205

Asn Asp Gly Ile Pro Asp Ser Leu Glu Val Glu Gly Tyr Thr Val Asp
    210                 215                 220

Val Lys Asn Lys Arg Thr Phe Leu Ser Pro Trp Ile Ser Asn Ile His
225                 230                 235                 240

Glu Lys Lys Gly Leu Thr Lys Tyr Lys Ser Ser Pro Glu Lys Trp Ser
                245                 250                 255

Thr Ala Ser Asp Pro Tyr Ser Asp Phe Glu Lys Val Thr Gly Arg Ile
            260                 265                 270

Asp Lys Asn Val Ser Pro Glu Ala Arg His Pro Leu Val Ala Ala Tyr
        275                 280                 285

Pro Ile Val His Val Asp Met Glu Asn Ile Ile Leu Ser Lys Asn Glu
    290                 295                 300

Asp Gln Ser Thr Gln Asn Thr Asp Ser Gln Thr Arg Thr Ile Ser Lys
```

305 310 315 320

Asn Thr Ser Thr Ser Arg Thr His Thr Ser Glu Val His Gly Asn Ala
325 330 335

Glu Val His Ala Ser Phe Phe Asp Ile Gly Gly Ser Val Ser Ala Gly
340 345 350

Phe Ser Asn Ser Asn Ser Ser Thr Val Ala Ile Asp His Ser Leu Ser
355 360 365

Leu Ala Gly Glu Arg Thr Trp Ala Glu Thr Met Gly Leu Asn Thr Ala
370 375 380

Asp Thr Ala Arg Leu Asn Ala Asn Ile Arg Tyr Val Asn Thr Gly Thr
385 390 395 400

Ala Pro Ile Tyr Asn Val Leu Pro Thr Thr Ser Leu Val Leu Gly Lys
405 410 415

Asn Gln Thr Leu Ala Thr Ile Lys Ala Lys Glu Asn Gln Leu Ser Gln
420 425 430

Ile Leu Ala Pro Asn Asn Tyr Tyr Pro Ser Lys Asn Leu Ala Pro Ile
435 440 445

Ala Leu Asn Ala Gln Asp Asp Phe Ser Ser Thr Pro Ile Thr Met Asn
450 455 460

Tyr Asn Gln Phe Leu Glu Leu Glu Lys Thr Lys Gln Leu Arg Leu Asp
465 470 475 480

Thr Asp Gln Val Tyr Gly Asn Ile Ala Thr Tyr Asn Phe Glu Asn Gly
485 490 495

Arg Val Arg Val Asp Thr Gly Ser Asn Trp Ser Glu Val Leu Pro Gln
500 505 510

Ile Gln Glu Thr Thr Ala Arg Ile Ile Phe Asn Gly Lys Asp Leu Asn
515 520 525

Leu Val Glu Arg Arg Ile Ala Ala Val Asn Pro Ser Asp Pro Leu Glu
530 535 540

Thr Thr Lys Pro Asp Met Thr Leu Lys Glu Ala Leu Lys Ile Ala Phe
545 550 555 560

Gly Phe Asn Glu Pro Asn Gly Asn Leu Gln Tyr Gln Gly Lys Asp Ile
565 570 575

Thr Glu Phe Asp Phe Asn Phe Asp Gln Gln Thr Ser Gln Asn Ile Lys
580 585 590

Asn Gln Leu Ala Glu Leu Asn Ala Thr Asn Ile Tyr Thr Val Leu Asp
595 600 605

Lys Ile Lys Leu Asn Ala Lys Met Asn Ile Leu Ile Arg Asp Lys Arg
610 615 620

Phe His Tyr Asp Arg Asn Asn Ile Ala Val Gly Ala Asp Glu Ser Val
625 630 635 640

Val Lys Glu Ala His Arg Glu Val Ile Asn Ser Ser Thr Glu Gly Leu
645 650 655

Leu Leu Asn Ile Asp Lys Asp Ile Arg Lys Ile Leu Ser Gly Tyr Ile
660 665 670

Val Glu Ile Glu Asp Thr Glu Gly Leu Lys Glu Val Ile Asn Asp Arg
675 680 685

Tyr Asp Met Leu Asn Ile Ser Ser Leu Arg Gln Asp Gly Lys Thr Phe
690 695 700

Ile Asp Phe Lys Lys Tyr Asn Asp Lys Leu Pro Leu Tyr Ile Ser Asn
705 710 715 720

Pro Asn Tyr Lys Val Asn Val Tyr Ala Val Thr Lys Glu Asn Thr Ile
725 730 735

Ile Asn Pro Ser Glu Asn Gly Asp Thr Ser Thr Asn Gly Ile Lys Lys
740 745 750

Ile Leu Ile Phe Ser Lys Lys Gly Tyr Glu Ile Gly
755 760

<210> 31

<211> 431

<212> PRT

<213> Protein sequence of Clostridium botulinum C2 toxin component 1

<400> 31

Met Pro Ile Ile Lys Glu Pro Ile Asp Phe Ile Asn Lys Pro Glu Ser
1 5 10 15

Glu Ala Gln Lys Trp Gly Lys Glu Glu Glu Lys Arg Trp Phe Thr Lys
20 25 30

```
Leu Asn Asn Leu Glu Glu Val Ala Val Asn Gln Leu Lys Thr Lys Glu
        35                  40                  45

Asp Lys Thr Lys Ile Asp Asn Phe Ser Thr Asp Ile Leu Phe Ser Ser
    50                  55                  60

Leu Thr Ala Ile Glu Ile Met Lys Glu Asp Glu Asn Gln Asn Leu Phe
65                  70                  75                  80

Asp Val Glu Arg Ile Arg Glu Ala Leu Leu Lys Asn Thr Leu Asp Arg
                85                  90                  95

Glu Val Ile Gly Tyr Val Asn Phe Thr Pro Lys Glu Leu Gly Ile Asn
            100                 105                 110

Phe Ser Ile Arg Asp Val Glu Leu Asn Arg Asp Ile Ser Asp Glu Ile
        115                 120                 125

Leu Asp Lys Val Arg Gln Gln Ile Ile Asn Gln Glu Tyr Thr Lys Phe
    130                 135                 140

Ser Phe Val Ser Leu Gly Leu Asn Asp Asn Ser Ile Asp Glu Ser Ile
145                 150                 155                 160

Pro Val Ile Val Lys Thr Arg Val Pro Thr Thr Phe Asn Tyr Gly Val
                165                 170                 175

Leu Asn Asn Lys Glu Thr Val Ser Leu Leu Leu Asn Gln Gly Phe Ser
            180                 185                 190

Ile Ile Pro Glu Ser Ala Ile Ile Thr Thr Ile Lys Gly Lys Asp Tyr
        195                 200                 205

Ile Leu Ile Glu Gly Ser Leu Ser Gln Glu Leu Asp Phe Tyr Asn Lys
    210                 215                 220

Gly Ser Glu Ala Trp Gly Glu Lys Asn Tyr Gly Asp Tyr Val Ser Lys
225                 230                 235                 240

Leu Ser Gln Glu Gln Leu Gly Ala Leu Glu Gly Tyr Leu His Ser Asp
                245                 250                 255

Tyr Lys Ala Ile Asn Ser Tyr Leu Arg Asn Asn Arg Val Pro Asn Asn
            260                 265                 270

Asp Glu Leu Asn Lys Lys Ile Glu Leu Ile Ser Ser Ala Leu Ser Val
        275                 280                 285

Lys Pro Ile Pro Glu Thr Leu Ile Ala Tyr Arg Arg Val Asp Gly Ile
```

290 295 300

Pro Phe Asp Leu Pro Ser Asp Phe Ser Phe Asp Lys Lys Glu Asn Gly
305 310 315 320

Glu Ile Ile Ala Asp Lys Thr Lys Leu Asn Glu Phe Ile Asp Lys Trp
325 330 335

Thr Gly Lys Glu Ile Glu Asn Leu Ser Phe Ser Ser Thr Ser Leu Lys
340 345 350

Ser Thr Pro Leu Ser Phe Ser Lys Ser Arg Phe Ile Phe Arg Leu Arg
355 360 365

Leu Ser Glu Gly Thr Ile Gly Ala Phe Ile Tyr Gly Phe Ser Gly Phe
370 375 380

Gln Asp Glu Gln Glu Ile Leu Leu Asn Lys Asn Ser Thr Phe Lys Ile
385 390 395 400

Phe Arg Ile Thr Pro Ile Thr Ser Ile Ile Asn Arg Val Thr Lys Met
405 410 415

Thr Gln Val Val Ile Asp Ala Glu Val Ile Gln Asn Lys Glu Ile
420 425 430

<210> 32

<211> 721

<212> PRT

<213> Protein sequence of Clostridium botulinum C2 toxin component 2

<400> 32

Met Leu Val Ser Lys Phe Glu Asn Ser Val Lys Asn Ser Asn Lys Asn
1 5 10 15

Tyr Phe Thr Ile Asn Gly Leu Met Gly Tyr Tyr Phe Glu Asn Asp Phe
20 25 30

Phe Asn Leu Asn Ile Ile Ser Pro Thr Leu Asp Gly Asn Leu Thr Phe
35 40 45

Ser Lys Glu Asp Ile Asn Ser Ile Leu Gly Asn Lys Ile Ile Lys Ser
50 55 60

Ala Arg Trp Ile Gly Leu Ile Lys Pro Ser Ile Thr Gly Glu Tyr Ile
65 70 75 80

```
Leu Ser Thr Asn Ser Pro Asn Cys Arg Val Glu Leu Asn Gly Glu Ile
                85                  90                  95

Phe Asn Leu Ser Leu Asn Thr Ser Asn Thr Val Asn Leu Ile Gln Gly
            100                 105                 110

Asn Val Tyr Asp Ile Arg Ile Glu Gln Leu Met Ser Glu Asn Gln Leu
        115                 120                 125

Leu Lys Asn Tyr Glu Gly Ile Lys Leu Tyr Trp Glu Thr Ser Asp Ile
    130                 135                 140

Ile Lys Glu Ile Ile Pro Ser Glu Val Leu Leu Lys Pro Asn Tyr Ser
145                 150                 155                 160

Asn Thr Asn Glu Lys Ser Lys Phe Ile Pro Asn Asn Thr Leu Phe Ser
                165                 170                 175

Asn Ala Lys Leu Lys Ala Asn Ala Asn Arg Asp Thr Asp Arg Asp Gly
            180                 185                 190

Ile Pro Asp Glu Trp Glu Ile Asn Gly Tyr Thr Val Met Asn Gln Lys
        195                 200                 205

Ala Val Ala Trp Asp Asp Lys Phe Ala Ala Asn Gly Tyr Lys Lys Tyr
    210                 215                 220

Val Ser Asn Pro Phe Lys Pro Cys Thr Ala Asn Asp Pro Tyr Thr Asp
225                 230                 235                 240

Phe Glu Lys Val Ser Gly Gln Ile Asp Pro Ser Val Ser Met Val Ala
                245                 250                 255

Arg Asp Pro Met Ile Ser Ala Tyr Pro Ile Val Gly Val Gln Met Glu
            260                 265                 270

Arg Leu Val Val Ser Lys Ser Glu Thr Ile Thr Gly Asp Ser Thr Lys
        275                 280                 285

Ser Met Ser Lys Ser Thr Ser His Ser Ser Thr Asn Ile Asn Thr Val
    290                 295                 300

Gly Ala Glu Val Ser Gly Ser Leu Gln Leu Ala Gly Gly Ile Phe Pro
305                 310                 315                 320

Val Phe Ser Met Ser Ala Ser Ala Asn Tyr Ser His Thr Trp Gln Asn
                325                 330                 335

Thr Ser Thr Val Asp Asp Thr Thr Gly Glu Ser Phe Ser Gln Gly Leu
            340                 345                 350
```

Ser Ile Asn Thr Gly Glu Ser Ala Tyr Ile Asn Pro Asn Ile Arg Tyr
355 360 365

Tyr Asn Thr Gly Thr Ala Pro Val Tyr Asn Val Thr Pro Thr Thr Thr
370 375 380

Ile Val Ile Asp Lys Gln Ser Val Ala Thr Ile Lys Gly Gln Glu Ser
385 390 395 400

Leu Ile Gly Asp Tyr Leu Asn Pro Gly Gly Thr Tyr Pro Ile Ile Gly
405 410 415

Glu Pro Pro Met Ala Leu Asn Thr Met Asp Gln Phe Ser Ser Arg Leu
420 425 430

Ile Pro Ile Asn Tyr Asn Gln Leu Lys Ser Ile Asp Asn Gly Gly Thr
435 440 445

Val Met Leu Ser Thr Ser Gln Phe Thr Gly Asn Phe Ala Lys Tyr Asn
450 455 460

Ser Asn Gly Asn Leu Val Thr Asp Gly Asn Asn Trp Gly Pro Tyr Leu
465 470 475 480

Gly Thr Ile Lys Ser Thr Thr Ala Ser Leu Thr Leu Ser Phe Ser Gly
485 490 495

Gln Thr Thr Gln Val Ala Val Val Ala Pro Asn Phe Ser Asp Pro Glu
500 505 510

Asp Lys Thr Pro Lys Leu Thr Leu Glu Gln Ala Leu Val Lys Ala Phe
515 520 525

Ala Leu Glu Lys Lys Asn Gly Lys Phe Tyr Phe His Gly Leu Glu Ile
530 535 540

Ser Lys Asn Glu Lys Ile Gln Val Phe Leu Asp Ser Asn Thr Asn Asn
545 550 555 560

Asp Phe Glu Asn Gln Leu Lys Asn Thr Ala Asp Lys Asp Ile Met His
565 570 575

Cys Ile Ile Lys Arg Asn Met Asn Ile Leu Val Lys Val Ile Thr Phe
580 585 590

Lys Glu Asn Ile Ser Ser Ile Asn Ile Ile Asn Asp Thr Asn Phe Gly
595 600 605

Val Gln Ser Met Thr Gly Leu Ser Asn Arg Ser Lys Gly Gln Asp Gly

610 615 620

Ile Tyr Arg Ala Ala Thr Thr Ala Phe Ser Phe Lys Ser Lys Glu Leu
625 630 635 640

Lys Tyr Pro Glu Gly Arg Tyr Arg Met Arg Phe Val Ile Gln Ser Tyr
645 650 655

Glu Pro Phe Thr Cys Asn Phe Lys Leu Phe Asn Asn Leu Ile Tyr Ser
660 665 670

Ser Ser Phe Asp Lys Gly Tyr Tyr Asp Glu Phe Phe Tyr Phe Tyr Tyr
675 680 685

Asn Gly Ser Lys Ser Phe Phe Asn Ile Ser Cys Asp Ile Ile Asn Ser
690 695 700

Ile Asn Arg Leu Ser Gly Val Phe Leu Ile Glu Leu Asp Lys Leu Ile
705 710 715 720

Ile